The Patent Office

INVESTOR IN PEOPLE

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77 bears an amendment, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears a correction, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed P. Mahoney

Dated 23 NOV 1999

Patents Form 1/77

Patents Act 1977
(Rule 16)

The Patent Office

1/77

THE PATENT OFFICE
31 JAN 1998
RECEIVED BY POST

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
Gwent NP9 1RH

1. Your reference

P57031P

02FEB98 E334601-1 D02813
P01/7700 25.00 - 9801990.4

2. Patent application number
(The Patent Office will fill in this part)

9801990.4

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

~~PIC~~
~~Fyfield Wick~~
~~Abingdon~~
~~Oxfordshire~~
~~OX13 5NA~~

Pig Improvement Company UK Limi
100 George Street
LONDON
W1S 5RH

Patents ADP number *(if you know it)*

~~UK~~ ~~AL 17/3/98~~ DEW 4/8/98

If the applicant is a corporate body, give the country/state of its incorporation

~~3700596~~ 7482672001

4. Title of the invention

METHODS FOR ANALYZING ANIMAL PRODUCTS

5. Name of your agent *(if you have one)*

~~Fry Heath & Spence~~

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

~~The Old College~~
~~53 High Street~~
~~Horley~~
~~Surrey RH6 7BN~~

KILBURN & STRODE
20 RED LION STREET
LONDON
WC1R 4PJ
125001

Patents ADP number *(if you know it)*

05880273001

51/77 27/3/98

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.
See note (d))

Yes

RECEIVED
FEB -7 2000
TECHNOLOGY CENTER 3700

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form	
Description	46
Claim(s)	5
Abstract	
Drawing(s)	14 + 14

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent *(Patents Form 7/77)*	
Request for preliminary examination and search *(Patents Form 9/77)*	
Request for substantive examination *(Patents Form 10/77)*	
Any other documents *(please specify)*	

11. I/We request the grant of a patent on the basis of this application.

Signature

Date 29th January 1998

12. Name and daytime telephone number of person to contact in the United Kingdom V A Price - Tel: 01293 776880

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered 'Yes' Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) For details of the fee and ways to pay please contact the Patent Office.

METHODS FOR ANALYZING ANIMAL PRODUCTS

The present invention relates to methods for analyzing animal products. In particular, the invention relates to methods for differentiating animal products on the basis of breed origin, determining or testing the breed origin of an animal product and for validating an animal product, as well as to kits for carrying out such methods.

Introduction

Animal breeds

For thousands of years, selective pressure has been applied by humans in the course of animal husbandry to produce livestock exhibiting certain desirable characteristics. These characteristics have been selected to meet aesthetic, technical, ritual, social and economic needs. The result has been the production of a large number of different animal *breeds*.

The term "breed" is a term of art used to define a homogenous, subspecific group of domestic livestock with definable and identifiable external characteristics that enable it to be separated by visual appraisal from other similarly defined groups within the same species. The term therefore defines a group of animals to which selective pressure has been applied by humans to give rise to a uniform appearance that is inheritable and distinctive with respect to other members of the species.

As breeds become established, their integrity is maintained by breed societies, herdbooks and pedigree records.

Breed selection

Conventional breed selection methods are based on direct measurement of the phenotype of an animal and/or its relatives. Thus, the implementation of breeding schemes requires extensive phenotypic record keeping. For example, dairy herd improvement programs in the United States and Western Europe relied in part on the collection of individual records (milk yield and composition, type traits, health traits, etc.) performed on a monthly basis for millions of cows. Likewise, breeding companies carefully monitor their pig and poultry breeding stock for a whole range of phenotypic measurements.

However, some important characteristics are not immediately apparent at the level of the living animal. For example, many parameters of meat quality are determined by subtle physiological or biochemical characteristics which are not readily apparent and so cannot serve as the basis for efficient artificial selection.

Breeding for qualities of this type has relied in part upon selection for other (more readily apparent traits) which are to some extent coinherited (linked or associated) with the desirable characteristics. For example, in the pig industry lop ears have in the past been associated with mothering ability and so have been used as a marker for this trait.

Conventional breed selection methods are limited by the fact that some phenotypes are expressed only in one sex or at a specific developmental stage. Moreover, some phenotypes are difficult and costly to measure. Indirect detection of such phenotypic traits *via* DNA-based diagnosis (for use in marker-assisted selection or MAS) is therefore seen as a desirable alternative to direct measurement of phenotypic parameters (see Georges and Andersson (1996), *Livestock genomics comes of age*, Genome Research, Vol. 6: 907-921). However, the gene structure-function relationships underlying many of the desirable traits are often highly complex and not yet sufficiently well-established to make such an approach feasible in practice.

Breed identification

The definition of animal breeds is currently at a watershed. Whereas previously they have been defined by overt physical characteristics and pedigree records, in the future as new breeds are developed from specific breeding lines they will be defined by sets of DNA markers. The work described herein allows not only the most accurate approach to breed determination currently possible in a range of products but also allows the integration in a common format of breed determinant information obtained through use of the present invention with that which will be used in the future. The present invention therefore allows not only the determination of source breed in the current environment but also links this to the development of future breeds and their unique identification.

It is generally recognized that the only definitive way to identify a particular animal as a representative of a given breed is through its pedigree. Thus, despite the fundamental importance of overt phenotypic traits in the breeding process and in the maintenance of breed purity, those skilled in the art generally consider that breed identity cannot be definitively characterized on the basis of visual inspection of such traits. By way of example, the genetic factor causing the belt phenotype in pigs is dominant to the non-belted form. Thus, a belted animal may result from an animal of a belted breed such as Hampshire being crossed with a non-belted breed.

As stated in *PIGS* A handbook to breeds of the world, V. porter, Helm Inf. Ltd, ISBN 1-873403-17-8, 1993, page 16, "What is a Breed?":

> *Appearances can be deceptive:* never *judge a pig breed by its coat!*

However, in many circumstances breed identification on the basis of direct evidence of pedigree is difficult or impossible. Thus, in practice, so-called "breed markers" may be used to determine breed identity.

The term "breed marker" is a term of art which defines a measurable characteristic which on the basis of empirical data appears to be breed specific. Breed markers include genotypic features such as DNA polymorphisms, chemical features such as protein and water contents of meats, epigenetic/biochemical features (such as protein polymorphisms), chromosome structure, gene copy number, DNA fingerprinting, microsatellite analysis and RAPD DNA markers.

Other useful markers include *breed determinants*. The term "breed determinant" is used herein to indicate an overt phenotypic characteristic which is used (at least in part) as the basis of artificial selection during breeding programmes. It is used in contradistinction to the term "breed marker", which (as explained above) is used herein to define other characteristics which appear to be breed specific on the basis of empirical data. The term "breed determinant gene" is used to indicate a gene which is involved (at least in part) in the expression of the corresponding overt phenotypic characteristic.

Some breed determinants (e.g. coat colour) have traditionally been used as breed "trademarks", and so have long served as an indication of pedigree (and breed identity). Other breed determinants that have also been selected for in breed development include features such as ear carriage, face shape and general anatomical conformation. The advantage of breed determinants relative to simple breed markers is the inseverable link between the characteristics of the breed and the determinant.

Biochemical and genetic tests for breed identity

Many of the breed markers discussed above can be characterized using biochemical or genetic tests. Such markers include genotypic features (e.g. DNA polymorphisms), biochemical features (e.g. protein polymorphisms), chromosome structure, gene copy number, DNA fingerprints, microsatellite patterns and RAPD DNA markers.

However, there are significant problems associated with such tests, as discussed below:

Tests based on the chemical composition of animal products (e.g. meat or seminal plasma) may be compromised by the fact that the chemical profile varies between muscles and is affected by diet, age, sex and sample storage conditions. Moreover, the results obtained are usually quantitative in nature, leading to problems with interpretation and comparison between different test sites.

Tests based on protein polymorphisms are limited by the fact that the distribution of any given protein is unlikely to be uniform, so that the protein of interest is absent in certain tissues. Thus, a number of different polymorphism markers may be required to check all products of interest for breed provenance. Moreover, such tests are based on antibody assays, and a significant investment is also required to develop the reagents for a specific antibody test.

Chromosome structure analyses are compromised by the high level of skill required for cytogenetical methodology and interpretation and the elaborate precautions and care required for sample preservation. Such markers are poorly applicable on anything but materials derived from living or newly deceased animals.

Classical DNA fingerprinting is based upon regions of repeated DNA sequence that due to their structure show a large degree of variation in length within a population. Such regions are often present in a number of copies within the DNA of an individual, thus increasing the potential for individual variation. By separating fragments of the total DNA according to size and then defining the position (and so size) of the hypervariable region using a specific probe, a fingerprint of a series of bands for a particular individual can be obtained. A number of probes for hypervariable regions of DNA have been examined in pigs (including M13 viral sequences and human minisatellite probes) and it is claimed that specific bands were found in each breed.

Random Amplified Polymorphic DNA (RAPD) markers are based upon PCR amplification of DNA fragments using primers of random sequence. Such reactions generally give rise to a number of DNA fragments which can be characterised according to size by gel electrophoresis. If the products of reactions based upon DNA from different breeds are examined there is the possibility of finding certain DNA bands which are breed specific. However, there is in most cases no direct link between the alleles of such repeat series present and the features determining the actual nature of the breed. This, combined with the hypervariable nature of these regions of DNA, results in them rarely being breed specific (similar alleles being found in a number of different breeds). As there is no link to the phenotype of the breed there is a greater risk that cross specific alleles could exist or arise in a breed, whereas this is unlikely with breed determinants as they define the phenotype itself. Given the large number of populations of animals of specific breeds that exist, extensive research would have to be carried out to exclude a DNA marker from breeds other than that with which it is claimed to be linked.

However, a major drawback with this approach is that RAPD markers are considered to be unreliable and found to be subject to variation between laboratories. Such problems are exacerbated when samples of different types and history must be analyzed and compared.

There is therefore a need for reliable breed markers which can be used as the basis for rapid and inexpensive methods for identifying the breed provenance of various animal products and for validating animal products (such as foodstuffs and semen for use in breeding programmes).

It has now been recognized that breed determinants as hereinbefore defined (such as coat colour) have unexpected advantages as breed identifiers or breed specific markers. In particular, it has surprisingly been discovered that the use of overt phenotypic characteristics as the basis for selection over long periods of time has led to particular alleles becoming fixed in most breeds. Such breed markers can be used to provide industry standard profiles for a particular breed that has application to all materials derived from a particular species.

Thus, it has now been found that many breeds are in fact genetically homogenous with respect to breed determinant genes (as hereinbefore defined), so that these genes may serve as the basis of reliable breed-specific markers (contrary to the prejudice in the art mentioned earlier regarding the utility of breed determinants *per se*, such as coat colour, in breed identification).

Moreover, it has surprisingly been found that the nature of the breed determinant genes (or alleles thereof) underlying any one breed determinant (such as coat colour) may be highly polymorphic. Thus, variation in breed determinant genes and/or alleles between different breeds may exist, notwithstanding the fact that the different breed determinant genes/alleles may contribute to the expression of the same overt phenotypic characteristic.

Prior to the present invention, it was assumed that the corresponding genetic determinants would be insufficiently polymorphic to provide a useful basis for distinguishing between breeds. For example, coat colour was known to be shared among different breeds of pig and (as mentioned above) was therefore not regarded as a good candidate for a breed specific marker. However, the present inventors have found that the alleles underlying the coat phenotype in such breeds are in fact highly polymorphic and often distinctive (and so useful as the basis for breed identification).

Similar considerations apply to other overt physical traits (breed determinants), which may therefore be shared by different breeds while nevertheless associated with distinct genes/alleles in each breed. An example of this is seen in cattle exhibiting the double muscled phenotype. Work by Kambadur *et alia* (1997, Genome Research 7, 910-915) and Grobet *et alia* (1997, Nature Genetics 17, 71-74) illustrates that the double muscled phenotype of cattle is caused by mutations in the myostatin gene. However, in the Belgian Blue and Asturiana breeds, this gene contains an 11 bp deletion whereas in the Piedmontese breed a G to A transition is present. Thus, as with porcine coat colour a single selected characteristic is caused by a number of potential polymorphisms. However, the nature of the

arisal and selection history for such overt physical characteristics leads to the fixation of particular alleles within the breeds contributing to the breed specific profile of determinants.

In the light of these findings, it has now been recognized that genetic analysis of breed determinants (such as coat colour) provides an effective means for validating animal products (e.g. foodstuffs) and may advantageously be incorporated into animal product (e.g. food) processing lines to monitor and maintain product quality and into quality control protocols in the food industry.

Coat colour

Coat colour is determined by the action of a number of different gene loci. For example, the gene determining whether a pig is white or coloured is designated *I* (for inhibition of coat colour). The version of the gene preventing the expression of any colour (*I*) is dominant to that which allows colour to develop (*i*). Traditional selection for white animals has reduced the frequency of *i*, but it still remains in the population of white heterozygous carrier animals. Recently, a number of structural differences in the alleles of the *KIT* gene were identified and found to be involved with this aspect of coat colour determination which allowed the development of methods of distinguishing between alleles at this locus.

However, animals which carry two copies of the recessive allele, *i*, at this locus have non-white coat colours (Johansson-Moller *et al.*, *Mamm. Genome*, 7:822-830 (1996), WO-A-97/05278, the disclosure of which is incorporated herein by reference). Pigs of this type can be all one colour, such as the Duroc (which is red), or have combinations of colours (particularly spotted or striped or banded patterns, such as the Pietrain and Hampshire, respectively). Many other combinations are possible and are observed (see the table, below):

Genotype	Colour
I/I	White
I/i	White
i/i	Coloured
Ip/Ip	White with coloured spots
Ip/I	White
Ip/i	White with coloured patches

The non-white colour in such animals may be varying shades of red or black. The type of colour expressed is determined by the action of a second gene which is designated *E* (for extension of coat colour).

The density and coverage of coat colour and the position of bands of white are determined by additional loci, which are not considered here.

The extension locus is known in other breeds of domestic animals, such as the horse, where *e* is associated with chestnut colour (Adalsteinsson, *J. Hered.* 65:15-20 (1974)), cattle (Klungland *et al.*, *Mammalian Genome* 6: 636-639 (1995)), the fox (Adalsteinsson, *J. Hered.* 78:15-20 (1987)) and the mouse (Jackson, *Ann Rev Genet.* 28: 189-217 (1994)). The extension locus encodes the alpha melanocyte-stimulating hormone receptor (αMSHR).

Classical segregation analyses have identified a minimum of three alleles at the pig extension locus: E for uniform black, E^p for black spotting and e for uniform red. The dominance relationship among the three alleles is as follows: $E > E^p > e$.

The present inventors have now found that these coat colour variations are associated with sequence polymorphism in the αMSHR in the pig. The DNA sequence of this gene has been analyzed using samples from the following breeds with different coat colour: Wild Boar which is wild type coloured, Meishan and Hampshire (which carry alleles for uniform black (E), Pietrain and Large White (which carry alleles for black spotting (E^p)) and Duroc (which is uniform red (e)).

In Large White the patches or spots of colour that might be expected due to the presence of the E^p allele are hidden as this breed also carries the dominant white gene which prevents any expression of colour. Four different αMSHR sequences were obtained from the breeds mentioned above. We have designated the allele found in the Wild Boar as E^+ and assume that the presence of this allele is necessary for the expression of the wild type colour.

The E alleles for uniform black carried by Meishan and Hampshire pigs were associated with different αMSHR sequences. We have denoted these two alleles E^m and E^h, respectively. The DNA sequence associated with the allele for black spotting (E^p) was unexpectedly indistinguishable from that of uniform black in Hampshire (E^h). The similarity of the E^p and E^h alleles suggests that they are derived from a common origin. The absence of a unique differentiating modification in the sequenced part of the gene implies that the difference is likely to be in the upstream or downstream part of the gene, possibly in a regulatory element of the genes.

Sequence comparisons of these regions from Pietrain and Hampshire or other breeds will provide information to develop kits specific for E^p and E^h. Alternatively, alleles of linked markers, such as microsatellite or AFLP markers, found to be in linkage disequilibrium with these alleles could be used to predict colour genotype. In conclusion, we have found four different αMSHR sequences associated

with five different extension alleles, i.e. E^+, E^m, E^h/E^p *and* e.

The colours of a series of pig breeds, the classical genotypes for *I* and *E* and the determined genotypes for *E* based on sequencing and testing studies are shown in the table below:

Breed	*I locus*	*E* locus
Hampshire	*i/i*	E^h/E^h
Large White	*I/I*	E^p/E^p
Landrace	*I/I*	E^p/E^p
Pietrain	*i/i*	E^p/E^p
Berkshire	*i/i*	E^p/E^p
Meishan	*i/i*	E^m/E^m
Duroc	*i/i*	*e/e*
Wild Boar	*i/i*	E^+/E^+

All of the DNA sequence differences identified in the αMSHR gene are single base pair changes. Some of these are silent while a number lead to changes in the amino acid sequence of the αMSHR protein. For example, the difference between *e* and the other alleles is a missense mutation in the coding sequence of the αMSHR gene. Importantly, the difference in the pig gene is different from that found in other species. The cattle and mouse *e* mutations are one base pair deletions, whilst the mutation identified here is a missense mutation in a region which is conserved among human, mouse, cattle and horse gene sequences. The E^p group has a single missense mutation which leads to a change in another conserved region, which is only four amino acid residues away from a modification observed in the dominant allele found in the fox. Finally, the Meishan allele (E^m) shows four nucleotide changes (two amino acid changes in the corresponding protein). Two of these differences are in the same region of the gene which is altered in cattle. Thus, it is possible to distinguish between the alleles of E^+, E^m, E^h/E^p and *e* and so determine the genotype of individual pigs (or the genetic provenance of products derived therefrom) with respect to non-white coat colour. Interestingly, the white breeds that have been examined all appear to be fixed for alleles E^p at the *E* locus. There is considered to be potentially some modifying effect of the *E* locus on the phenotype conferred by the *I* locus. While the basis of this is not established, the fixing of E^p in these lines illustrates the subtle effects on loci involved with coat colour upon selection for breed characteristics thus providing more determinants among such loci than might be expected.

Associations can be determined between extension locus genotype and microsatellite sequences which are linked to the gene. A number of microsatellite markers have been located to the region of porcine chromosome 6 to which the αMSHR gene has been mapped.

Summary of the invention

According to the present invention there is provided a method for differentiating animal products on the basis of breed origin, for determining or testing the breed origin of an animal product or for validating an animal product, wherein the method comprises the steps of: (i) providing a sample of the animal product; and (ii) analyzing the allele(s) of one or more breed determinant genes present in the sample.

As explained above, the breed determinant is an overt phenotypic trait. As used herein, an overt phenotypic trait is one which can be visually recognized.

Differentiation of animal products on the basis of breed origin involves the partition of members of a class of different animal products into a number of different products sharing the same breed origin. It does not necessarily imply identification of the nature of the breed source. Animal product differentiation of this kind basis may be sufficient where the consistency of source of animal products must be monitored (but its actual breed provenance is not important).

In contrast, *determination* of the breed origin of an animal product implies identification of the breed source, while *testing* the breed origin implies analysis sufficient to determine whether a breed source other than that desired has been used (without necessarily identifying such other breed sources in cases where they are indicated).

Validating an animal product implies confirming that it meet stipulated specifications as to breed provenance. Such validation may involve differentiation, determination and/or testing, depending on the circumstances under which the analysis is performed and the nature and extent of ancillary data which may be available.

The sample for use in the invention may be in any convenient form. In many cases, the sample will be a sample of a food (e.g. meat product). For most applications, the sample is pre-treated (e.g. extracted, purified and/or fractionated) in such a way so as to make the alleles of a breed determinant gene or genes available for analysis (either at the level of nucleic acids (such as RNA or DNA) and/or proteins). The sample is preferably a nucleic acid sample, in which case the analysing step (ii) comprises DNA or RNA analysis. Alternatively, the sample may be a protein sample (where the

nature of the protein reflects a breed determinant allele), in which case the analysing step (ii) comprises protein analysis.

The breed determinant of the invention may be a monogenic or polygenic trait. Monogenic traits are preferred, since the genes conferring such traits are relatively easily identified and analyzed. However, in some cases it may be useful to analyze the alleles of polygenic traits (i.e. traits which are controlled by a plurality of genes), since such the underlying allele polymorphism is often greater in such cases (so increasing the potential for breed differentiation).

Typically, overt phenotypic traits are those traits which have been used as the basis for artificial selection during the breeding programme. The overt phenotypic trait is preferably a behavioural or morphological, physiological or behavioural trait.

The overt phenotypic trait may vary qualitatively or quantitatively between breeds. Preferred are traits which vary qualitatively between breeds, since such traits are often reflected by qualitative differences in the alleles of the corresponding breed determinant gene(s). In such cases, analysis yields relatively robust positive-negative results, which are easily interpreted and compared between testing stations/laboratories.

The breed determinant gene analysed in step (ii) may be any suitable breed determinant gene. Such genes may be identified and analysed by methods well known in the art using routine trial and error. Preferably, they are selected from any of a coat colour, pattern, texture, density or length gene; a ear aspect gene; a double muscling gene; a horn morphology gene; a tusk morphology gene; an eye colour gene; a plumage gene; a beak colour/morphology gene; a vocalization (e.g. barking) gene; a comb or wattle gene; and/or a gene controlling display behaviour.

In preferred embodiments, the breed determinant gene is the KIT and/or αMSHR coat colour gene (for example, the pig KIT and/or αMSHR gene).

The analysis step (ii) may comprise any of a wide range of known nucleic acid/protein analytical techniques. The nature of the analytical technique selected is not critical to the practice of the invention, and those skilled in the art can readily determine the appropriate technique according to the circumstances in which the analysis is to be conducted and the type of data required.

Preferably, the analysis step (ii) comprises selectively amplifying a specific fragment of nucleic acid (e.g. by PCR), testing for the presence of one or more restriction endonuclease sites within the breed determinant gene(s) (e.g. restriction fragment length polymorphism (RFLP) analysis), determining the

nucleotide sequence of all or a portion of the breed determinant gene(s), probing the nucleic acid sample with an allele-specific DNA or RNA probe, or carrying out one or more PCR amplification cycles of the nucleic acid sample using at least one pair of suitable primers and then carrying out RFLP analysis on the amplified nucleic acid so obtained.

Alternatively, the analysis step (ii) comprises probing the protein sample with an antibody (e.g. a monoclonal antibody) specific for an allele-specific epitope, electrophoretic analysis, chromatographic analysis, amino-acid sequence analysis, proteolytic cleavage analysis or epitope mapping.

In particularly preferred embodiments, the analysis step (ii) comprises determining the nucleotide sequence of the KIT and/or αMSHR gene or the amino acid sequence of the KIT and/or αMSHR protein. Here, the analysis may comprise establishing the presence or absence of at least one mutation in the KIT and/or αMSHR gene. Any method for identifying the presence of the specific sequence change may be used, including for example single-strand conformation polymorphism (SSCP) analysis, ligase chain reaction, mutagenically separated PCR, RFLP analysis, heteroduplex analysis, denaturing gradient gel electrophoresis, temperature gradient electrophoresis, DNA sequence analysis and non-gel based systems such as TaqMan™ (Perkin-Elmer).

In the TaqMan™ system, oligonucleotide PCR primers are designed that flank the mutation in question and allow PCR amplification of the region. A third oligonucleotide probe is then designed to hybridize to the region containing the base subject to change between different alleles of the gene. This probe is labelled with fluorescent dyes at both the 5' and 3' ends. These dyes are chosen such that while in this proximity to each other the flourescence of one of them is quenched by the other and cannot be detected. Extension by Taq DNA polymerase from the PCR primer positioned 5' on the template relative to the probe leads to the cleavage of the dye attached to the 5' end of the annealed probe through the 5' nuclease activity of the Taq DNA polymerase. This removes the quenching effect allowing detection of the florescence from the dye at the 3' end of the probe. The discrimination between different DNA sequences arises through the fact that if the hybridization of the probe to the template molecule is not complete (i.e. there is mismatch of some form), then cleavage of the dye does not take place. Thus only if the nucleotide sequence of the oligonucleotide probe is completely complementary to the template molecule to which it is bound will quenching be removed. A reaction mix can contain two different probe sequences each designed against different alleles that might be present thus allowing the detection of both alleles in one reaction.

Although the TaqMan™ system is currently capable of distinguishing only two alleles, labelled probe primer sets could be developed in which the probes for certain target allele(s) are labelled with a different fluorescent dye from non target alleles. For example, if one wished to confirm that a group

of Duroc breed pigs carried only allele *e* one could have a probe present capable of detecting this allele labelled with one fluorescent dye and probes capable of detecting all the other alleles labelled with the second dye. Thus one would detect the presence of any non Duroc type alleles at this locus. Such probe sets could be designed and labelled according to the needs of the experiment.

The analysis step (ii) may further comprise determining the association between one or more microsatellite marker alleles linked to the KIT and/or αMSHR gene and to particular alleles of the KIT and/or αMSHR gene.

Alternatively, the analysis step (ii) may be based on the identification of microsatellite markers present in the nucleic acid sample.

The analysis step (ii) preferably comprises: (a) determining the association between one or more microsatellite marker alleles linked to the KIT and/or αMSHR gene and to particular alleles of the KIT and/or αMSHR gene; determining which microsatellite marker allele or alleles are present in the nucleic acid sample.

The analysis step (ii) preferably further comprises the step of determining the genotype of at least one additional locus, for example an additional breed determinant (e.g. coat colour) locus. Particularly preferred as an additional locus is the *KIT* gene locus (e.g. the pig *KIT* gene locus).

The analysis step (ii) preferably comprises PCR using at least one pair of suitable primers. In the case where the gene is the pig αMSHR gene, the at least one pair of suitable primers is:

αMSHR Forward Primer 1: (5'-TGT AAA ACG ACG GCC AGT RGT GCC TGG AGG TGT-3') or
αMSHR Reverse Primer 5: (5'-CGC CCA GAT GGC CGC GAT GGA CCG-3')
αMSHR Forward Primer 2: (5'-CGG CCA TCT GGG CGG GCA GCG TGC-3') or
αMSHR Reverse Primer 2: (5'-GGA AGG CGT AGA TGA GGG GGT CCA-3')
αMSHR Forward Primer 3: (5'-GCA CAT CGC CCG GCT CCA CAA GAC-3') or
αMSHR Reverse Primer 3: (5'-GGG GCA GAG GAC GAC GAG GGA GAG-3')

The analysis step (ii) may also comprise restriction fragment length polymorphism (RFLP) analysis, for example involving digesting the pig nucleic acid with one or more of the restriction enzymes *Bst*UI, *Hha*I and/or *Bsp*HI. In cases where the gene is the pig αMSHR gene, this analysis may involve identification of a polymorphism at nucleotide position 400, 422, 480, 487, 608, 844, 846 or in the region 482-487 or 841-849 of the coding sequence of the pig αMSHR gene.

The analysis step (ii) may involve carrying out one or more PCR amplification cycles of the nucleic acid sample using at least one pair of suitable primers and then carrying out RFLP analysis on the amplified nucleic acid so obtained to determine the KIT or αMSHR genotype of the pig. Here, when the gene is the pig αMSHR gene the at least one pair of suitable primers is as defined above.

The animal product preferably comprises or consists of meat (e.g. processed and/or canned meat), egg, egg swab or washing, semen, blood, serum, sputum, wool, biopsy sample or leather. It may comprise genomic DNA, RNA or mitochondrial DNA.

The animal is preferably a mammal (e.g. pig, cattle, dog, cat, horse, sheep, rodent or rabbit), fish (e.g. salmon or trout) or bird (e.g. chicken or turkey).

The invention may be used with extremely small samples and can be used to screen large numbers of samples quickly and inexpensively. The invention may be adapted to yield absolute results, and quantification is not essential. Moreover, only small fragments of nucleic acid are required, and the same tests can be used on the majority of animal products.

Applications

The invention finds application in a number of areas. For example, certain breeds are considered to yield meat of higher eating quality, and a number of retailers now market products which claim to be derived from specific or traditional breeds (for example, Wild Boar crosses). The invention enables consumer organisations to validate these claims and also permits retailers to monitor the quality of the products with which they are being supplied (i.e. perform product validation). The invention finds particular application in validation studies carried out and used by retailers to support consumer confidence, since the linkage between a genetic marker and an overt physical feature is more readily grasped by the lay person than the concept of breed specific markers. This makes the use of such breed determinants attractive and also offers marketing opportunities for retailers to underpin validation schemes.

There are also a number of reports of breed influences on the quality of hams produced by various meat processing techniques. For example, in one report hams from three different pig breeds were reliably classified on the basis of sensory descriptors of marbling, saltiness and dry cure flavour. The breed identification processes of the invention enables producers to validate raw materials as part of quality control.

The ability to enforce and validate raw material source uniformity also yields improved process control, lower costs and greater product consistency, since it has now been found that heterogeneity in chemical

composition of products from different breeds is an important factor in flavour profile variation and there may also be differences in the functionality of other meat components between breeds.

The invention also finds utility in the maintenance of stock purity by animal (e.g. pig) breeders. The small size of traditional breed populations means that the maintenance of a gene pool of sufficient size to avoid the effects of inbreeding requires the importation and movement of stock between separate populations. A risk of genetic contamination is associated with such movements, and the invention may be used to reduce or eliminate these risks. The maintenance of biodiversity and the rare breeds providing the reservoir for this diversity provides an increasing need for breed identifiers.

The invention may also be used as part of a breeding programme to confirm particular crosses. This may be of enormous value in the establishment of pyramid breeding schemes. Particular breed characteristics such as coat colour, body shape and ear aspect are often altered in such crosses, yet there is a need to be able to confirm the presence of genetics of the desired parents.

Such visible breed characteristics for the visible confirmation of crosses are also absent in the use of artificial insemination, where semen may be supplied from pigs in distant geographical locations.

Brief description of the Figures

Figure 1: Partial nucleotide sequence (a) and the derived amino acid sequence (b) of the porcine αMSHR gene as determined from a number of pig breeds. Position numbers are in accordance with the bovine BDF3 sequence (Vanetti *et alia*, *FEBS Lett.*, 348: 268-272 (1995).

Figure 2: Agarose gel electrophoresis of DNA fragments obtained by digestion of DNA fragments amplified from the porcine αMSH-R gene with *Bst*UI or *Hha*I. Lanes labelled M contain DNA markers of 50, 150, 300,500, 750, 1000bp. The other samples were derived from:

1. Pietrain
2. Pietrain
3. Large White
4. Large White
5. Large White
6. Duroc
7. Duroc
8. Hampshire
9. Meishan
10. Berkshire
11. Berkshire

Figure 3: Agarose gel electrophoresis of DNA fragments obtained by digestion of DNA fragments amplified from the porcine αMSH-R gene with *Bst*UI (lanes labelled B) or *Hha*I (lanes labelled H). Lanes labelled M contain DNA markers of 50, 150, 300,500, 750, 1000bp. The other samples were derived from:

1. Retailer 1. Skin
2. Retailer 1. Fat
3. Retailer 1. Muscle
4. Retailer 2. Fat
5. Retailer 2. Muscle

Figure 4: Electropherogram (4% agarose) showing RT-PCR products of KIT exon 16-19 with the primers KIT1F and KIT7R. The samples 1-3 and 4-6 are Swedish Large White and Hampshire pigs respectively. The size difference between the 424 and 301 bp fragments is due to lack of exon 17 in the latter fraction. The two upper bands of the Yorkshire pigs were interpreted as heteroduplexes (HD).

Figure 5: A 48 bp sequence is shown comprising 21 bp of KIT exon 17 and 27 bp of KIT intron 17. The position of the intron/exon border is marked with a vertical line and the splice site mutation (nt1 $^{G\rightarrow A}$) indicated with a vertical arrow. Identical bases in alleles *I*^{P} and *i* are marked with a dot.

Figure 6: *Nla*III PCR RFLP test used to detect the presence of a splice site mutation in intron 17 of the KIT gene. Figure 6A shows the position of two *Nla*III recognition sites within the PCR product amplified using primer pair KIT21 and KIT35. All distances are given in base pairs. Figure 6B shows the size of fragments which result following *Nla*III digestion of either normal *KIT* or splice mutant *KIT*. Figure 6C illustrates use of the PCR RFLP test. Lane 1 shows the KIT21/KIT35 amplified fragment undigested. Digestion was performed on PCR products amplified from, in Lane 2: a clone which contains the splice site mutation; Lane 3: a clone which contains the normal splice site sequence; Lane 4: genomic DNA from a coloured pig; Lane 5: genomic DNA from a white pig. Fragment sizes are given in base pairs.

Figure 7: Comparison of the ratio of normal to splice mutant KIT for three classes of genotype.

Figure 8: Comparison of the ratio of normal to splice mutant KIT for two breeds of pig.

Figure 9: SSCP analysis of the KIT gene in Swedish Landrace (lanes 1-8) and Wild Boar (lanes 9 & 10) breeds. The two polymorphic bands are indicated.

Figure 10: Nucleotide sequence of the porcine *KIT* cDNA from an animal of the Hampshire breed. The sequence is numbered with the first nucleotide of the N terminal methionine codon taken as 1.

Figure 11: Polyacrylamide gel electrophoresis of PCR-RFLP analysis of KIT gene at polymorphic nucleotide 2678 in a number of animals. Lanes: 1 & 2, Hampshire Wild Boar respectively, both homozygous for the C at position 2678. Lanes 3-7 and 9 & 10, unrelated Large White sows all homozygous for T at position 2678. Lane 11, a Pietrain, homozygous for T at this position and lane 8 a Large White sow heterozygous for C and T. Lane 12 contains undigested PCR product and lane M DNA size standards.

Examples

Example 1: Determination of the sequence of the αMSHR gene

The DNA sequence of the porcine αMSHR gene was determined through the DNA sequencing of a combination of PCR products and cloned portions of porcine DNA.

Preparation of template DNA for PCR

DNA can be prepared from any source of tissue containing cell nuclei, for example white blood cells, hair follicles, ear notches and muscle. The procedure here relates to blood cell preparations; other tissues can be processed similarly by directly suspending material in K buffer and then proceeding from the same stage of the blood procedure. The method outlined here produces a cell lysate containing crude DNA which is suitable for PCR amplification. However, any method for preparing purified, or crude, DNA should be equally effective.

Blood was collected in 50mM EDTA pH 8.0 to prevent coagulation. 50μl of blood was dispensed into a small microcentrifuge tube (0.5ml Eppendorf or equivalent). 450μl of TE buffer was added to lyse the red blood cells (haem groups inhibit PCR) and the mix vortexed for 2 seconds. The intact white and residual red blood cells were then centrifuged for 12 seconds at 13,000 g in a microcentrifuge. The supernatant was removed by gentle aspiration using a low pressure vacuum pump system. A further 450μl of TE buffer was then added to lyse the remaining red blood cells and the white blood cells collected by centrifugation as before. If any redness remained in the pellet, this process was repeated until the pellet was white. After removal of the last drop of supernatant from the pelleted white blood cells, 100μl of K buffer containing proteinase K was added and the mixture incubated at 55 degrees C for 2 hours. The mixture was then heated to 95-100 degrees C for 8 minutes and the DNA lysates stored at -20 degrees C until needed.

Reagents

T.E. Buffer:	10mM TRIS-HCl pH8.0
	1mM EDTA
K Buffer:	50mM KCl
	10mM TRIS-HCl pH8.3
	2.5mM MgCl2
	0.5% Tween 20

PCR to produce DNA sequencing template

The αMSHR gene was amplified for sequence analysis using three primer pairs.

Primers MSHR Forward Primer 1: (5'-TGT AAA ACG ACG GCC AGT RGT GCC TGG AGG TGT CCA T-3'); and

MSHR Forward Primer 5: (5'-CGC CCA GAT GGC CGC GAT GGA CCG-3')

amplify a 428 bp fragment from the 5' half of the gene.

Primers MSHR Forward Primer 2: (5'-CGG CCA TCT GGG CGG GCA GCG TGC-3');

and αMSHR Reverse Primer 2: (5'-GGA AGG CGT AGA TGA GGG GGT CCA-3')

amplify a 405 bp fragment the 3' half of the gene.

As these two fragments are non-overlapping a third primer pair
αMSHR Forward Primer 4 (5'-TGC GCT ACC ACA GCA TCG TGA CCC TGC-3'); and

αMSHR Reverse Primer 4 (5'-GTA GTA GGC GAT GAA GAG CGT GCT-3')

were used to amplify a 98 bp fragment which spans the 50 bp gap. PCR was carried out on a DNA thermal cycler (Perkin Elmer 9600) in a total volume of 20 μl containing 25 ng genomic DNA, 1.0 mM MgCl2, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 200 (M dNTPs, 0.5 U AmpliTaq Gold (Perkin Elmer) and 10 pmol of both forward and reverse primer. To activate AmpliTaq Gold, initial heat denaturation was carried out at 94 degrees C for 10 minutes followed by 32 cycles each consisting of 45 sec at 94 degrees C, 45 sec at 53 degrees C and 45 sec at 72 degrees C. The final extension lasted for 7 min at 72 degrees C. PCR products were cloned into vector pUC18 using the

SureClone ligation kit (Pharmacia).

Preparation of plasmid DNA

Plasmid DNA was purified from overnight bacterial culture using the Jetstar plasmid midi kit 50 (Genomed) and the resulting DNA diluted to 150 ng/μl.

Sequencing of plasmid DNA

Cloned plasmid inserts were sequenced using dye primer chemistry. Each cycling reaction was prepared with template and ready reaction mix containing fluorescently labelled M13 forward or reverse primer as described in the ABI Prism protocol P/N 402113 (Perkin Elmer). Cycling and sample pooling was performed using a Catalyst 800 Molecular Biology Workstation (ABI) following the instruments user manual (Document number 903877, Perkin Elmer). The resulting extension products were purified, loaded and analysed using the 377 ABI Prism DNA sequencer as described by the instrument protocol (Perkin Elmer protocol P/N 402078).

Dye Terminator Sequencing of PCR products

Dye terminator DNA sequencing requires purification of PCR product free from excess dNTPs and residual primers. This was achieved by passage of the template DNA through QiaQuick spin columns (Qiagen) before the purified DNA was diluted to 15 ng/μl. Dye terminator cycle sequencing was performed using AmpliTaq DNA polymerase FS in accordance with the ABI Prism protocol P/N 402078 (Perkin Elmer). Cycle sequencing reactions were performed in a total reaction volume of 10 μl. This comprised 1.6 pmole of either the forward or reverse primer used to amplify the target fragment from genomic DNA, 20 ng of purified template DNA and terminator ready reaction mix (Perkin Elmer) which contains each of four dye terminators, dNTPs, Tris-HCl (pH 9.0), $MgCl_2$, thermal stable pyrophosphate and AmpliTaq DNA polymerase FS. Cycle sequencing was performed with a GeneAmp 9600 machine (Perkin Elmer) over 25 cycles, each consisting of 10 sec at 96 degrees C, 5 sec at 50 degrees C and 4 min at 60 degrees C. Extension products were purified for gel separation using ethanol precipitation, loaded and run on a 377 ABI Prism DNA sequencer as described by the instrument protocol (Perkin Elmer protocol P/N 402078).

Results

The partial coding region DNA sequence of the porcine αMSHR gene sequence from a number of pig breeds is given in figure 1a. The derived amino acid sequence is shown in figure 1b.

Example 2: PCR-RFLP based discrimination of alleles at the *E* locus

DNA preparation for PCR

As in example 1.

PCR

Reactions were set up in a 20μl reaction volume in thin walled 0.25ml tubes (Perkin Elmer) with the following components:

20μl reaction volume:
2μl template DNA
1.5 mM $MgCl_2$
200μM each dNTP,
3pM each of forward and reverse primers
0.5 U AmpliTaq Gold (Perkin Elmer)

MSHR Forward primer 3 sequence: 5' GCA CAT CGC CCG GCT CCA CAA GAC 3'
MSHR Reverse primer 3 sequence: 5' GGG GCA GAG GAC GAC GAG GGA GAG 3'

The reaction tubes were placed on a Perkin Elmer 9600 thermal cycler preheated to 94 degrees C and PCR carried out according to the regime below: -

Initial denaturation step of 94 degrees C for 10 min.

33 cycles: 94 degrees C - 45 secs
53 degrees C - 45 secs
72 degrees C - 45 secs

The last cycle is followed by a final elongation of 72 degrees C for 7 min. Samples are stored at 4 degrees C until required.

Restriction Enzyme Digestion and Electrophoresis

The PCR amplification product is 148bp in length. To test for polymorphism in the amplified products the reaction is split into two aliquots of 10μl each of which is digested with *Hha*I (GIBCO-BRL) or *Bst*UI (New England Biolabs). The reactions are set up and incubated as below:

***Bst*UI digest**	***Hha*I digest**
10μl amplified DNA	10μl amplified DNA
2.5μ *Bst*UI	2.5μ *Hha*I
60 degrees C 60 minutes	0.5μl 10x React 2 buffer (GIBCO-BRL)
	37 degrees C 60 minutes

Following digestion, 2μl of loading dye is added to each reaction (100mM Tris pH8.0, 100mM

Boric Acid, 1mM EDTA, 50% (v/v) glycerol, 0.02% w/v Orange G) and the mixes loaded on a 4% agarose gel (3% NuSieve/1% Seakem, FMC Bioproducts) in 0.5x TBE (44.5mM Tris pH8.0, 44.5 mM boric acid and 0.5mM EDTA) and electrophoresed for 1 hour at 150v.

Products are visualised by ethidium bromide staining.

Results

*Bst*UI and *Hha*I digestion each result in bands of 61 and 87bp. The relationship of digestion to the possible allele is as shown in the table below:

Relationship of restriction digest profiles to individual alleles at the *E* locus

Allele	Digestion with *Bst*UI	Digestion with *Hha*I
$E^+/E^p/E^h$	Yes	Yes
E^m	No	Yes
e	No	No

If the uncut alleles are designated as allele 1 and the alleles digesting with each enzyme as allele 2 the various genotypes will be as shown in the table below:

Actual E genotypes and associated scores

Genotype	*Bst*UI	*Hha*I
E^m/E^m	1/1	2/2
E^m/E^p or E^m/E^+	1/2	2/2
E^m/e	1/1	1/2
E^p/E^p or E^p/E^+ or E^+/E^+	2/2	2/2
E^p/e or E^+/e	1/2	1/2
e/e	1/1	1/1

Note: The results for animals carrying the allele E^h will be the same as those carrying E^p.

Samples were prepared from a number of pigs and tested according to the above protocol. The results are shown in the table below and figure 2 illustrates the patterns seen upon electrophoresis.

E genotypes determined for a range of breeds using the *Bst*UI/*Hha*I digestion system

Breed	No Tested	Genotype (see note 1)	αMSHR type	
			*Bst*UI	*Hha*I
Hampshire	9	E^P/E^P	2/2	2/2
Large White	4	E^P/E^P	2/2	2/2
Landrace	1	E^P/E^P	2/2	2/2
Pietrain	3	E^P/E^P	2/2	2/2
Berkshire	2	E^P/E^P	2/2	2/2
Bazna	4	E^P/E^m	1/2	2/2
		E^m/E^m	1/1	2/2
Duroc	4	*e/e*	1/1	1/1
Meishan	3	E^m/E^m	1/1	2/2

Note 1. Where the genotype E^P is listed this cannot be distinguished from E^+ or E^h.

As can be seen from the results above the genotypes determined fit with those expected from the sequencing data given in figure 1a for Hampshire, Large White, Meishan and Duroc. The additional breeds typed here show the genotypes expected from their phenotype and descriptions in published literature (Ollivier and Sellier, Ann. Génét. Sél. Anim., 14: 481-544, (1982)). The Pietrain is a white breed with black patches of varying extent and has long been considered to be E^P (in agreement with the result here). The Berkshire, originally a spotted breed, is now a mainly black animal with white 'socks' again generally considered to be E^P as was found here. The Landrace is a white animal due to it carrying the dominant white allele at the *I* locus, however its genotype at the *E* locus has been shown to be E^P from classical breeding studies. Once again this is in agreement with the results obtained here. The Bazna is a Romanian breed having black base colour with a white belt. It was developed from the Berkshire and Mangalitza, a Hungarian breed with a number of colour variations including black (Porter, Pigs, a handbook to breeds of the world, publ: Helm Information, ISBN 1-873403-17-8 (1993)). The ancestory of the Bazna being based upon a black breed potentially carrying a similar allele to the Meishan, E^m, and the Berkshire carrying E^P, is in agreement with the alleles found to be present in the breed in this work.

Example 3: Validation of source breeds of retail meats

DNA preparation

DNA was prepared from different parts of pork chops from two separate retailers. The DNA was prepared from skin (1 retailer only), fat and muscle using the Promega Wizard Genomic DNA

preparation kit according to the manufacturers instructions. Approximately 4mm^3 of each tissue was cut into small fragments for the extraction.

PCR and Restriction Digest Analysis

This was carried out exactly as in example 2.

Results

The results are shown in figure 3. It can be seen that DNA extracted from a range of tissue types can be utilised for this DNA based test with results being obtained here for muscle, fat and skin. The genotype of the pig with regard to the (MSHR gene can then be determined. In this case the material from both retailers was derived from an animal of test type *Bst*UI 1/2 and *Hha*I 1/2 using the nomenclature as in example 2. This translates into genotype E^p/e or E^+/e. Based on our current knowledge of the distribution of the alleles in commercial pig breeds the conclusion can be drawn that both source animals contain genetic material derived from the Duroc.

Example 4: Validation of source breeds of processed meat samples

Method

DNA was prepared from heat treated meat samples according to the method of Meyer *et al.* (Journal of AOAC International, 78 1542-1551). Meat samples were minced with a scalpel and 0.3g transferred to a sterile 1.5ml eppendorf tube containing 430μl of extraction buffer (10mM Tris-HCl pH 8.0, 150mM NaCl, 2mM EDTA, and 1% w/v sodium dodecyl sulphate). Fifty microlitres of 5M guanidine hydrochloride and 20μl of 20mg/ml proteinase K (Boehringer) were added and mixed by inversion followed by incubation at 57°C for 3h. After digestion samples were centrifuged for 10 min at 13,000 x g, and 450μl of the aqueous phase added to 1ml Wizard DNA purification resin (Promega). The mixture was mixed by gentle inversion and following the Wizard DNA clean-up procedure carried out according to the manufacturers instructions the purified DNA was eluted with 50μl of 70°C water. 1μl of a 1:10 dilution was then used as template in a 10μl PCR.

PCR was carried out as described in the previous example.

Results

Meat samples from a Large White based line and a Duroc based line heated at 80°C for 30 mins could be differentiated on the basis of their genotype at the *E* locus with the Large White samples giving a pattern characteristic of the E^p allele and the Duroc samples a pattern characteristic of the *e* allele.

Example 5: Validation of source breeds of semen

Genomic DNA was isolated from porcine semen. 1ml of semen was centrifuged for 2 min at 13,500 x g and the supernatant removed. 1 ml of 2xSSC was added and the mix vortexed to resuspend the sperm. The mix was then centrifuged as before and the supernatant removed. 400μl of 0.2M NaOAc pH 7.0 was added and the mix vortexed followed by the addition of 34μl of ß-mercaptoethanol. The mixture was incubated at 40°C for 30 min followed by the addition of 100μl of 10% w/v sodium dodecyl sulphate and 50μl of 15 mg/ml Proteinase K (Boehringer) and further incubation at 40°C for 3 hours. 500μl phenol equilibrated with Tris-HCl pH 8.0 was added and the mix vortexed twice followed by centrifugation at 13,500 x g for 4 min. 400μl of the aqueous phase was removed and 800μl of ethanol added. DNA was allowed to precipitate for 5 min at room temperature followed by centrifugation at 13,500 x g for 5 min. The pellet was washed with 800μl 70% ethanol v/v and air dried followed by resuspension in 200μl of Wizard DNA resuspension buffer (Promega). 1μl of a 1/10 dilution was used in a 10μl PCR

PCR was carried out as described in example 2.

Results

Semen form a Hampshire based line and a Duroc based line could be differentiated on the basis of their genotype at the *E* locus with the Hampshire samples giving a pattern characteristic of the E^n allele and the Duroc samples a pattern characteristic of the *e* allele.

Example 6: Discrimination of allele E^+ from alleles E^n/E^n

DNA preparation

DNA was prepared as described in example 1.

PCR

Reactions were set up in a 20μl reaction volume in thin walled 0.25ml tubes (Perkin Elmer) with the following components:

10μl reaction volume:

- 2μl template DNA
- 2.5 mM $MgCl_2$
- 200μM each dNTP.
- 5pmol each of forward and reverse primers
- 0.5 U AmpliTaq Gold (Perkin Elmer)

Forward primer sequence: 5' CTG CCT GGC CGT GTC GGA CCT G 3'

Reverse primer sequence: 5' CTG TGG TAG CGC AGC GCG TAG AAG 3'

The reaction tubes were placed on a Strategene Robocycler and PCR carried out according to the regime below: -

Initial denaturation step of 94°C for 10 min.

30 cycles: 94°C - 60 secs
61°C - 60 secs
72°C - 60 secs

The last cycle is followed by a final elongation of 72°C for 7 min. Samples are held at 6°C until required.

Restriction Enzyme Digestion and Electrophoresis

The PCR amplification product is 228 in length. To test for polymorphism in the amplified products the reaction is digested with *BspHI* (New England Biolabs). The reactions are set up and incubated as below:

***BspHI* digest**
10μl amplified DNA
1ul 10x React 2 (NEB New England Biolabs)
0.5μl deionised water
5 units *Bst*UI
37°C 60 minutes

Following digestion, 2μl of loading dye is added to the reaction (100mM Tris pH8.0, 100mM Boric Acid, 1mM EDTA, 50% (v/v) glycerol, 0.02% w/v Orange G) and the mix loaded on a 4% agarose gel (3% NuSieve/1% Seakem, FMC Bioproducts) in 0.5x TBE (44.5mM Tris pH8.0, 44.5 mM boric acid and 0.5mM EDTA) and electrophoresed for 1 hours at 150v.

Products are visualised by ethidium bromide staining.

Results

BspHI digestion each result in bands of 124 and 104bp. The relationship of digestion to the possible allele is as shown below:

Relationship of restriction digest profiles to individual alleles at the *E* locus

Allele	Digestion with *BspHI*
E^n/E^p	Yes
E^+	No

Samples were prepared from a number of pigs and tested according to the above protocol and the results are shown below:

E genotypes determined for a range of breeds using the *BspHI* digestion system

Breed	No Tested	Genotype (see note 1)	Number
Wild Boar x Swedish Landrace	3	E^p/E^+	3
Large White	4	E^p/E^p	4
Landrace	1	E^p/E^p	1
Pietrain	3	E^p/E^p	3

Note 1. Where the genotype E^p is listed this cannot be distinguished from E^n

Example 7: Discrimination of cattle products by breed

DNA was prepared from cattle muscle samples as described in example 4. PCR was then carried out in a 100μl reaction using the primer pair:

5'-TGAGGTAGGAGAGTTTTGGG-3' and
5'-TCGAAATTGAGGGGAAGACC-3'

as described in Kambadur *et al. Genome Research* 7: 910-915 (1997) at a concentration of 500nM with other reaction components being 2.5mM $MgCl_2$, 200μM dNTPs, 50mM KCl, 10mM Tris-HCl pH 8.3, 5 units AmpliTaq Gold (Perkin Elmer). 1μl of bovine genomic DNA was used as template. Denaturation was carried out for 12 min at 94°C followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, 72° 1.5 min followed by 5 min at 72°C. Following PCR 2.0μl of loading dye (44.5mM Tris pH 8.0, 44.5mM boric acid, 0.5mM EDTA, 50%w/v glycerol, 0.02% w/v Orange G) was added to 10μl of product and analysis carried out by electrophoresis on a 2% agarose gel prepared in 0.5x TBE buffer (44.5mM Tris pH 8.0, 44.5mM boric acid, 0.5mM EDTA) for 1 hour at 100V.

The remainder of the PCR was analysed for DNA sequencing using ABI dye terminator chemistry as described in example 1.

Results

Bovine myostatin DNA polymorphisms and related phenotype

Breed	Phenotype	nt position 941	length PCR product (bp)
Belgian Blue	Double muscle	G	482
Piedmontese	normal	A	493
Holstein-Friesian	Double muscle	G	493

Example 8

RT-PCR of porcine *KIT* exon 16-19

i. mRNA purification from blood samples

Fresh blood samples were collected in citrate tubes from coloured Hampshire pigs and Large White pigs. Leukocytes were isolated from 5 ml blood using Ficoll 100 (Pharmacia Biotech). Isolation of mRNA from leukocytes was then carried out using the Quickprep *Micro* mRNA purification kit (Pharmacia Biotech). The mRNA was stored as a precipitate under ethanol at -70°C for up to one month before use in reverse transcriptase (RT)-PCR.

ii.RT-PCR of KIT exon 16-19

First-strand cDNA synthesis was accomplished using the First-Strand cDNA Synthesis kit (Pharmacia Biotech) so that ~100 ng mRNA was randomly primed by 0.1 μg pd(N6) in a total volume of 15 μl. Two μl of the completed first cDNA strand reaction was then directly used per 12 μl PCR reaction by adding 10 μl PCR mix containing 10 pmol each of the mouse/human derived primers KIT1F and KIT7R (5'-TCR TAC ATA GAA AGA GAY GTG ACT C and 5'-AGC CTT CCT TGA TCA TCT TGT AG, respectively; Moller et al. 1996, *supra*), 1.2 μl 10 x PCR-buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl) and 0.5 U of AmpliTaq polymerase (Perkin-Elmer) incubated with an equal amount Taqstart antibody (Clontech) at 25°C for 5 min to achieve a hot start PCR. The reaction was covered with 20 μl mineral oil and thermocycled in a Hybaid Touchdown machine (Hybaid) with 40 cycles at 94°C for 1 min, 55-48°C (touchdown one degree per cycle the first seven cycles and then 48°C in the remaining cycles) for 1 min and 72°C for 1 min. After PCR 2μl loading dye was added to each sample which were then loaded on 4% agarose gel (Nusieve/Seakem 3:1, FMC Bioproducts) and electrophoresed with 100V for 80 min. Products were visualised by ethidium bromide staining and UV-illumination.

iii. Cloning and sequencing of RT-PCR-products

The RT-PCR products representing *KIT* exon 16-19 were purified by extraction from 2% agarose gels using the QIAEX gel extraction kit (QIAGEN) and cloned into the pUC18 vector using the Sureclone ligation kit (Pharmacia Biotech). Plasmids were isolated using the QIAFilter plasmid Midi kit (QIAGEN). Cloned plasmid inserts were sequenced using dye primer chemistry. Each cycling reaction was prepared with plasmid template DNA and ready reaction mix containing fluorescently labelled M13 forward or reverse primer as described in the ABI Prism protocol P/N 402113 (Perkin Elmer). Cycling and sample pooling were performed using a Catalyst 800 Molecular Biology Workstation (ABI) following the instruments user manual (Document number 903877, Perkin Elmer). The resulting extension products were purified, loaded and analysed using the 377 ABI Prism sequencer as described by the instrument protocol P/N 402078 (Perkin Elmer).

iv Results and discussion

A 424 bp fragment including KIT cDNA exon 16-19 was amplified from all pigs. The Hampshire pigs did not show any additional products whereas the Large White pigs (eight tested) all showed a 301 bp truncated cDNA fragment (Fig 4). Sequence analysis revealed the 424 bp fragment was identical in the two breeds whereas the whole exon 17 (123 bp) was missing from the 301 bp fragment. Apparent differences between individuals regarding the relative amounts of these two products may have been caused either by different genotypes containing differing numbers of copies of the *KIT* gene sequence, individual differences in mRNA expression levels or random RT-PCR effects.

The two upper fragments present in Large white pigs represent heteroduplexes between the 301 and 424 bp fragments (Fig. 2). This was shown by an experiment where these slow migrating fragments were generated by pooling homoduplexes of the 424 and 301 bp which were then heat denatured and cooled to 25°C. Moreover, cloning of the lower heteroduplex fraction of a Large White pig resulted in clones with insert length corresponding to either of the two homoduplexes.

Example 9

PCR Amplification and Sequencing of KIT Exon 17-Intron 17 (5' Splice Site)

i. PCR to produce DNA Sequencing Template

A 175 bp region including the boundary between exon17 and intron17 of the *KIT* gene was amplified for sequence analysis using forward primer KIT21 (5' - GTA TTC ACA GAG ACT TGG CGG C -3') and reverse primer KIT35 (5' - AAA CCT GCA AGG AAA ATC CTT CAC GG - 3'). PCR was carried out on a DNA thermal cycler (Perkin Elmer 9600) in a total volume of

20 μl containing 25 ng genomic DNA, 1.0 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl, pH 8.3, 200 μM dNTPs, 0.5 U AmpliTaq Gold (Perkin Elmer) and 10 pmol of both KIT21 and KIT35 primer. To activate AmpliTaq Gold, initial heat denaturation was carried out at 94°C for 10 minutes followed by 32 cycles each consisting of 45 sec at 94°C, 45 sec at 55°C and 45 sec at 72°C. The final extension lasted for 7 min at 72°C. PCR products were cloned into vector pUC18 using the SureClone ligation kit (Pharmacia Biotech).

ii. Preparation of Plasmid DNA

Plasmid DNA was purified from overnight bacterial culture using the Jetstar plasmid midi kit (Genomed) and the resulting DNA diluted to 150 ng/μl.

iii. Sequencing of plasmid DNA

DNA was sequenced as in example 8.

iv. Results

A portion of the DNA sequence from exon 17 and intron 17 of the KIT gene was determined and compared between animals with each of these three alleles. Figure 5 shows that the *I* allele carries a splice site mutation at position 1 of intron 17. This G to A base substitution is present in one of the two gene copies carried on each chromosome. The base substitution occurs in the invariant GT dinucleotide which characterises 5' exon/intron boundaries. Analysis of the I^P allele showed the splice site mutation was not present in either the normal (KIT1) or duplicated copy of the gene (KIT2). We have found the splice site mutation is unique to the *I* alleles, and therefore makes it possible to distinguish the I-KIT2 sequences.

Example 10

Testing For the Presence of the Splice Site Mutation with PCR RFLP

To easily test for the presence of the G to A splice site mutation, restriction endonuclease *Nla*III (CATG) was used to exploit the point substitution identified at position 1 of intron 17 (Figure 5). The *Nla*III recognition sites in the fragment amplified from *KIT* and the expected restriction products are illustrated in Figure 6A and 6B respectively.

i. PCR to produce DNA for RFLP Test

The PCR to produce DNA for RFLP analysis was performed exactly as described in example 9.

ii. Restriction Enzyme Digestion and Electrophoresis

The PCR amplification product is 175 bp in length. To test for polymorphism at position 1 of intron 17, digestion reactions were set up as below:

3.0 μl PCR amplified DNA
1.0 μl 10 X NEBuffer 4
0.1 μl BSA 100 μg/ml
0.1 μl *Nla*III 10 U/μl
5.8 μl dH2O

(1 X NEBuffer 4 (New England Biolabs) contains 50 mM potassium acetate, 20 mM Tris acetate, 10 mM magnesium acetate and 1 mM DTT). Following incubation at 37°C for 90 minutes each 10 μl reaction volume had 2 μl of loading dye added and the mix loaded on a 8% native polyacrylamide gel (Protogel, 37.5:1 acrylamide:bisacrylamide , National Diagnostics, Atlanta) in 0.5 X TBE (44.5 mM Tris pH 8.0, 44.5 mM boric acid and 0.5 mM EDTA) and electrophoresed for 3 hours at 200V in a vertical slab unit (SE600 Hoefer Scientific Instruments). Products were visualised by ethidium bromide staining.

iii Results

A PCR RFLP protocol was designed to test for the presence of the splice site mutation as the substitution occurs within the recognition site for restriction endonuclease *Nla*III. Figure 6B illustrates that presence of the G to A base substitution at position 1 of KIT intron 17 results in restriction at each of two *Nla*III recognition sites within the 175 bp DNA fragment. Following electrophoresis, this results in fragments of sizes 80 bp, 54 bp and 41 bp. Where the splice site mutation is absent however, incubation with *Nla*III results in digestion only at recognition site 1. Following electrophoresis this results in fragments of 134 bp and 41 bp. The invariant *Nla*III recognition site 1 serves as an internal control to ensure complete digestion has taken place. Results of this PCR RFLP analysis are illustrated in Figure 6C. Analysis was performed on fragments amplified from clones which either carry the splice site mutation (lane 2) or carry the normal splice site sequence (lane 3). Lane 4 shows the result of analysis where DNA amplified from the genomic DNA of a coloured animal was used. Lane 5 shows the resulting bands where a white animal was tested. The test was used to analyse 121 individuals from seven different breeds of pig. The splice site mutation was found only in the 97 animals with the dominant white phenotype (*I*/- or *I**/*i*) and none of the 24 coloured (I^p or *i*) examples (see table below). This analysis confirms *I* and *I** to be unique in that they are the only alleles to carry the splice site mutation.

Distribution of the Splice Site Mutation Between Different Breeds and Coat Phenotype

Breed	Coat Colour	Assumed Genotype[1]	Animals Tested	Normally spliced KIT[2]	Splice Mutation[2]
Large White	white	I/-	33	33	33
Landrace	white	I/-	56	56	56
Hampshire	coloured	i/i	5	5	0
Duroc	coloured	i/i	5	5	0
Pietrain	coloured	i/i	8	8	0
Meishan	coloured	i/i	5	5	0
Wild Boar	coloured	i/i	1	1	0
Wild Boar x Large White	white	I*/-	8	8	8
Totals					
	white	I/-	89	89	89
	white	I*/-	8	8	8
	coloured	i/i	24	24	0

[1] White animals may be homozygous or heterozygous for the *I* allele
[2] Presence of the splice site mutation determined by *Nla*III PCR RFLP test

Example 11

Quantification of Normal KIT and Splice Mutant KIT (Intron 17 nt1 $^{G\rightarrow A}$)

As the splice site mutation is present in only one of the duplicated regions of *I* and not in the duplicated region of I^P, the various genotypes can be expected to have the attributes described in the table below:

Genotype	Copies of Normal KIT	Copies of KIT containing the splice mutation	Ratio of normal KIT to splice mutant KIT
I/I	2	2	1:1
I/i	2	1	2:1
i/i	2	0	2:0
I/I^P	3	1	3:1
I^P /i	3	0	3:0

Due to the dominance of allele *I*, three of the genotypes in the table are carried by white animals and

therefore can not be identified by phenotypic characterisation. Quantification of the relative amounts of the normal *KIT* gene and the splice mutant *KIT* gene allows the ratio between the two to be calculated, and therefore the genotype of individual animals predicted. This was achieved by quantification of two DNA fragments following *Nla*III digestion. The amount of 134 bp fragment, representative of the normally spliced *KIT* gene, and of 54 bp fragment, representative of the splice mutant *KIT*, were measured following electrophoresis using GeneScan software.

i. PCR to Produce DNA for Quantification

As described in example 9 section i. The reverse primer KIT35 is labelled with the ABI fluorescent dye FAM at the 5' end.

ii Restriction Enzyme Digestion

As described in example 9 section ii.

iii Electrophoresis and Quantification of DNA Fragments

Following digestion, 0.5 μl of the reaction volume was mixed with 2.5 μl of deionised formamide, 0.5 μl of GS350 DNA standard (ABI) and 0.4 μl blue dextran solution before being heated to 90°C for 2 minutes and rapidly cooled on ice. Three μl of this mix was then loaded onto a 377 ABI Prism sequencer and the DNA fragments separated on a 6% polyacrylamide gel in 1 X TBE buffer for 2 hours at 700 V, 40 mA, 32 W. The peak area of fragments representative to both the normal and splice mutant forms of *KIT* were quantitated using the GeneScan (ABI) software.

iv. Ratio Calculations

The peak area value of the 134 bp fragment (normal *KIT*) was divided by twice the peak area value of the 54 bp fragment (splice mutant *KIT*) in order to calculate the ratio value for each sample.

v. Results

Analysis was performed on animals from the Swedish wild pig/Large White intercross pedigree for which genotypes at *I* have been determined by conventional breeding experiments with linked markers. Figure 7 and the table below show the ratio of normal to mutant *KIT* calculated for animals from each of the three genotype classes, *I/I* (expected ratio 1:1), *I/i* (expected ratio 2:1) and I/I^p (expected ratio 3:1). The results are entirely consistent with the expected ratio values and indicate that the three genotype classes can be distinguished using this method.

Ratio of the Two KIT Forms in Different Dominant White Genotypes in a Wild Pig/Large White Intercross

Genotype	Phenotype	Expected Ratio (Normal: Mutant)	Observed Ratio (Normal:Mutant) ± SE	Number Tested
I/I	white	1:1	1.15 ± 0.075	13
I/I^P	white	3:1	3.11 ± 0.084	12
I/i	white	2:1	2.23 ± 0.109	14

Figure 7 illustrates that the range of ratio values calculated for the two genotypes *I/I* and I/I^P do not overlap. This enables animals carrying the I^P allele to be identified and the frequency of the allele within different pig breeds determined. Ratio values were calculated for 56 Landrace and 33 Large White animals and the results are shown in Figure 8. A clearly bimodal distribution is observed with 7 Landrace and 3 Large White individuals having a ratio value of approximately 3 or above, suggesting them to be heterozygous carriers for the I^P allele (genotype I/I^P). This means I^P has gene frequency estimates of 6.25% (7/112 chromosomes tested) and 4.5% (3/66 chromosomes tested) within the Landrace and Large White breeds respectively.

EXAMPLE 12

(i) DNA Preparation

DNA can be prepared from any source of tissue containing cell nuclei, for example white blood cells, hair follicles, ear notches and muscle. The procedure outlined here relates to blood cell preparations; other tissues can be processed similarly by directly suspending material in K buffer and then proceeding from the same stage of the blood procedure. The method outlined here produces a cell lysate containing crude DNA which is suitable for PCR amplification. However, any method for preparing purified, or crude, DNA should be equally effective.

Blood was collected in 50 mM EDTA pH 8.0 to prevent coagulation. 50 μl of blood was dispersed into a small microcentrifuge tube (0.5 ml Eppendorf or equivalent). 450 μl of TE buffer was added to lyse the red blood cells (haem groups inhibit PCR) and the mix vortexed for 2 seconds. The intact white and residual red blood cells were then centrifuged for 12 seconds at 13,000 g in a microcentrifuge. The supernatant was removed by gentle aspiration using a low pressure vacuum pump system. A further 450 μl of TE buffer was then added to lyse the remaining red blood cells and the white blood cells collected by centrifugation as before. If any redness remained in the

pellet, this process was repeated until the pellet was white. After removal of the last drop of supernatant from the pelleted white blood cells, 100 μl of K buffer containing proteinase K was added and the mixture incubated at 55°C for 2 hours. The mixture was then heated to 95-100°C for 8 minutes and the DNA lysates stored at -20°C until needed.

Reagents

TE buffer:	10 mM TRIS-HCl pH 8.0
	1 mM EDTA
K buffer:	50 mM KCl
	10 mM TRIS-HCl pH 8.3
	2.5 mM $MgCl_2$
	0.5% Tween 20

Prior to use for lysates, 10 μl of 20 mg/ml proteinase K (Molecular Probes Inc.) per 1.0 ml of K buffer was added.

(ii) PCR

Reactions were set up as follows in thin walled 0.25 ml tubes (Perkin Elmer):

4.0 μl 5 μM *CRC* Forward primer;
4.0 μl 5 μM *CRC* Reverse primer;
4.0 μl 5 μM *KIT*1-REV primer;
4.0 μl 5 μM *KIT*1-FOR primer;
4.0 μl 2 mM dNTPs (Pharmacia);
4.0 μl 35 mM $MgCl_2$.

A wax bead (PCR Gem 50, Perkin Elmer) was added and the tube placed in a Perkin Elmer 9600 thermal cycler. The tube was then raised to 80°C for 15 seconds followed by cooling to 4°C. A second set of reagents was then added to each tube as below:-

4.0 μl 10x buffer;
9.6 μl sterile deionised water;
0.4 μl (0.5 units) AmpliTaq DNA polymerase (Perkin Elmer);
2 μl DNA lysate.

Reaction tubes were then placed on a Perkin Elmer 9600 thermal cycler preheated to 94°C and PCR

carried out according to the regime indicated below:-

94°C for 4 minutes;
20 cycles of 94°C for 30 secs, 62°C for 30 secs and 72°C for 30 secs;
0°C until required.

The number of cycles may vary depending upon the tissue used as the DNA source.

KIT primers

Forward	GAATATTGTTGCTATGGTGATCTCC *KIT*1-FOR
Reverse	CCGCTTCTGCGTGATCTTCCTG *KIT*1-REV

CRC primers

Forward	CTGGATGTCCTGTGTTCCCTGT CRC-FORWARD
Reverse	AGGTTTGTCTGCAGCAGAAGCTC CRC-REVERSE

The reverse *KIT* primer and the forward *CRC* primer are labelled with the ABI fluorescent dye FAM at the 5' end.

(iii) Electrophoresis and Quantitation of DNA Fragments

1 μl of the PCR was mixed with 2.5 μl of deionised formamide, 0.5 μl of GS350 DNA standards, 0.4 μl blue dextran solution, heated at 90°C for 2 minutes followed by rapid cooling on ice. 3 μl of this mix were then loaded onto an AB1373 DNA sequencer and DNA fragments separated on a 6% polyacrylamide gel in 1 x TBE buffer for 2 hours at 700 V, 40 mA, 32 W. The fragments corresponding to the products from the *KIT* and *CRC* genes were quantitated using GeneScan software, the peak area for each of the bands being determined.

(iv) Results

The data given in the table below represents the results obtained from an experiment in which DNA lysates were produced from each of 23 animals, with two PCR tests being carried out on each lysate. The ratio of *KIT* peak area to *CRC* peak area was calculated for each PCR and the average taken of those samples from the same animal.

Animal	Genotype	KIT/CRC peak area ratio
1	II	3.25
2	Ii	2.45
3	II	2.94
4	ii	1.16
5	ii	1.87
6	ii	1.20
7	Ii	2.18
8	Ii	2.19
9	II	2.88
10	ii	1.30
11	Ii	1.84
12	II	2.84
13	ii	1.50
14	ii	1.30
15	Ii	2.09
16	ii	1.31
17	ii	1.14
18	Ii	2.08
19	Ii	1.87
20	Ii	2.00
21	ii	0.99
22	ii	1.15
23	II	2.80

The upper and lower limits for the ratio values from animals of the different genotypes *II, Ii* and *ii* in this experiment are as below:

Genotype	Upper Limit	Lower Limit
I/I	3.25	2.80
I/i	2.45	1.84
i/i	1.50	0.99

These results illustrate differentiation of the genotypes using this test.

EXAMPLE 13

The second test utilises unique sequences of DNA that are present at one end of the duplication (or both ends if the duplicated region is reversed relative to the rest of the gene or if the duplicated region does not occur in direct tandem with the non-duplicated region). Oligonucleotide primers for use in PCR are designed such that at the annealing temperatures used in the PCR process, they will anneal only to the junction regions at the end of the duplicated region. A PCR is then carried out using two pairs of oligonucleotides. One pair consists of the aforementioned primer spanning the junction region and a second primer a suitable distance away which allows amplification to occur only from *I* allele containing duplication. The second pair of primers allow amplification of a sequence present only as a single copy in the haploid genome. The product of this reaction, carried out in the same tube, functions as an internal standard as in the previous test. The ratio of product from the reaction specific to the junction region is measured relative to that from the single copy control sequence.

In this test there is a larger difference between the predicted ratios of the products from the different genotypes. The relative levels of product and their ratios are illustrated below:-

Genotype	Junction Product	Control Product	Ratio
II	2	2	1:1
Ii	1	2	1:2
ii	0	2	0:2

These larger ratios allow greater differentiation between the ranges of results obtained from the different genotypes, reducing risks of miss-scoring animals.

EXAMPLE 14

(i) DNA Preparation

DNA can be prepared as described in example 12.

(ii) PCR

Reactions were set up as follows in thin walled 0.25 ml tubes (Perkin Elmer):

2.0 µl 5 mM *KIT*DEL2-FOR primer;

2.0 µl 5 mM *KIT*DEL2-REV primer;

1.0 μl 2 mM dNTPs (Pharmacia);

1.2 μl 25 mM MgCl2

2.0 μl 10x buffer (without MgCl2)

0.1 μl (0.5 units) AmpliTaq DNA polymerase (Perkin Elmer);

2.0 μl DNA lysate;

9.7 μl sterile deionised water.

Reaction tubes were then placed on a Perkin Elmer 9600 thermal cycler and PCR carried out according to the regime indicated below:-

95°C for 1 minute;
3 cycles of 95°C for 15 secs, 50°C for 20 secs and 72°C for 40 secs;
27 cycles of 94°C for 15 secs, 50°C for 20 secs and 72°C for 50 secs;
72°C for 5 minutes;
4°C until required.

The number of cycles may vary depending upon the tissue used as the DNA source.

KIT primers

Forward	GAAAGTGA(C/T)GTCTGGTCCTAT(C/G)GGAT *KIT*DEL2-FOR
Reverse	AGCCTTCCTTGATCATCTTGTAG *KIT*DEL2-REV

(iii) Electrophoresis

1 μl of the PCR product was mixed with 3 μl loading buffer (95% deionised formamide, 10mM NaOH, 20mM EDTA, 0.05% bromophenolblue, 0.05% Xylene-cyanol), heated to 95°C for 3 minutes followed by rapid cooling on ice. The sample was then loaded on an 8% native polyacrylamide gel (Protogel, 37.5:1 Acrylamide:bisacrylamide, National Diagnostics, Atlanta) in 1 x TBE buffer (89mM Tris, 89mM boric acid, 2mM EDTA.Na2). The DNA fragments were separated by electrophoresis for 4.5 hours at 6W with a constant temperature of 20°C and 0.6 x TBE as running buffer in a vertical slab unit (SE600 Hoefer Scientific Instruments, San Francisco).

(iv) Visualisation of DNA fragments by silver staining

After electrophoresis the gel was incubated, with gentle agitation, in the fix solution for 20 minutes or until the tracking dyes were no longer visible. The gel was rinsed three times (2 minutes each with agitation) in deionised water. The gel was then incubated in the staining

solution for 40 minutes, with gentle agitation, followed by a brief wash (5-10 seconds) in deionised water and direct transfer to the developing solution. The gel was incubated in the developing solution until bands were clearly visible and then the development was terminated by adding an equal volume of fix solution. Finally, the gel was rinsed for 2 minutes in deionised water.

Reagents

Fix solution:	10% glacial acetic acid in deionised water
Staining solution:	2 g silver nitrate ($AgNO_3$) 3 ml 37% formaldehyde 2 litres deionised water
Developing solution:	60 g sodium carbonate (Na_2CO_3) dissolved in 2 liters deionised water. Immediately before use add 3 ml 37% formaldehyde and 400 ml sodium thiosulfate (10 mg/ml). The solution should be at a temperature of 10-12°C when used.

(v) Results

This SSCP analysis reveals an informative polymorphism so far only found in animals with the dominant white phenotype (Fig. 9). In lanes 1 to 8 the analysis was carried out on DNA from Swedish Landrace pigs carrying the dominant white colour and in lanes 9 and 10 DNA was from wild pigs of wild type colour. The polymorphic bands are indicated. The polymorphism is characterised by two unique fragments only present in animals carrying a duplicated *KIT* gene of allele type *I*. The fragments represent heteroduplexes of DNA strands from PCR products of unequal length representing the duplicated and non-duplicated copy of the *KIT* gene. The results of a screening test with this marker using 40 unrelated animals representing five breeds and 190 F2 animals from a Large White/Wild pig intercross are presented in the table below:

BREED	COLOUR	NO. OF ANIMALS	HETERODUPLEX	
			PRESENT	NOT PRESENT
SWEDISH LANDRACE	WHITE	10	10	0
SWEDISH LARGE WHITE	WHITE	8	8	0
SWEDISH HAMPSHIRE	COLOURED	10	0	10
SWEDISH DUROC	COLOURED	10	0	10
WILD PIG	COLOURED	2	0	2
LARGE WHITE/ WILD PIG INTERCROSS	WHITE PATCH COLOURED	131 9 50	106 0 0	25 9 50

The results show that this particular polymorphism is very closely associated with the presence of the *KIT* duplication. It is not completely associated with the duplication as some white animals did not show the heteroduplex pattern. The polymorphism is therefore an example of a closely linked genetic marker which by itself or in combination with other linked markers can be used to differentiate genotypes as regards the dominant white coat colour.

EXAMPLE 15

i) DNA extraction

DNA was prepared as in example 12.

ii) PCR

Reactions were set up in 0.25ml thin walled reaction tubes (Perkin Elmer) as follows:

0.5μl 5 μM *KIT*DEL1-FOR primer

0.5μl 5 μM *KIT*DEL1-REV primer

1.0μl 2mM dNTPs (Pharmacia)

1.0μl 15mM $MgCl_2$

1.0μl 10X buffer

4.9μl Sterile distilled water

0.1μl AmpliTaq DNA polymerase

1.0μl DNA lysate

Reaction tubes were then placed in a Perkin Elmer 9600 thermal cycler and PCR carried out according to the regime

94°C for 4 minutes;
21 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec
72°C for 4 min;
4°C until required.

The number of cycles used may vary depending on the tissue used as the source of the DNA.

Primers

forward	TGTGGGAGCTCTTCTCTTTAGG *KIT*DEL1-FOR
reverse	CCAGCAGGACAATGGGAACATCT *KIT*DEL1-REV

The reverse primer was labeled with the ABI fluorescent dye FAM at the 5' end.

iii) Electrophoresis and quantitation of DNA fragments

1μl of the PCR was mixed with 1.5μl of deionised formamide, 0.25μl of GS350 DNA standards, 0.25μl loading buffer (50mg/ml blue dextran, 25mM EDTA) and heated at 90°C for two minutes followed by rapid cooling on ice. 1.75μl of this was then loaded onto an ABI 377DNA sequencer and DNA fragments separated on a 4.12% polyacrylamide gel in 1x TBE buffer for two hours at 3000V, 60mA, 200W and 48°C. The 97bp and 93bp fragments corresponding to the products from the *KIT* gene template lacking the deletion and containing the deletion respectively were quantitated using GeneScan software, the peak area for each of the bands being determined.

Results

The data given in the table below represents the results obtained from an experiment in which DNA lysates were produced from each of 20 animals of known genotype with one PCR test being carried out on each lysate. The ratio of the peak area of the product from the DNA template not containing the four base pair deletion to that containing the deletion was calculated.

ANIMAL	GENOTYPE	Non del/del peak area ratio
1	*II*	1.347
2	*II*	1.21
3	*II*	1.33
4	*II*	2.267
5	*II*	0.444
6	*II*	0.444
7	*II*	8.387
8	*II*	0.994
9	*II*	1.673
10	*Ii*	1.056
11	*Ii*	1.751
12	*Ii*	1.83
13	*Ii*	1.83
14	*Ii*	0.631
15	*Ii*	1.975
16	*Ii*	2.147
17	*Ii*	1.901
18	*Ii*	1.749
19	*Ii*	2.103
20	*Ii*	2.026

For this small sample the value of 1.5 which is midway between the predicted ratio values for each genotype (expected ratio=2 for *Ii* and 1 for *II*) might be used as the dividing line for scoring the animals to either genotype. It can be determined from the table that 7/10 *II* and 9/10 *Ii* are identified as the correct genotype.

Example 16

Sequencing of KIT cDNA clones

mRNA was isolated from peripheral blood leukocytes from white (Landrace/Large White) and coloured (Hampshire) pigs using the Message Maker mRNA isolation system (Gibco BRL) with one mRNA selection from total RNA. 100ng poly(A)$^+$ mRNA was reverse-transcribed with random primers (First-Strand cDNA Synthesis kit, Pharmacia Biotech) and the product was used at a 1:10

dilution for RT-PCR using the proof-reading Advantage KlenTaq Polymerase (Clontech) according to the manufacturer's recommendation. The following primers were used to amplify almost the entire coding sequence and some of the 5' untranslated region: KIT40 (5'-GGC TCT GGG GGC TCG GCT TTG C) corresponding to the 5'untranslated region and KIT22S (5'- TCA GAC ATC TTC GTG GAC AAG CAG AGG) corresponding to exon 21; both primers had been designed using consensus sequence of the human and mouse *KIT* sequences in the GENBANK database. The RT-PCR products were gel purified and cloned using the pGEM-T vector system (Promega). Plasmid clones were sequenced using a set of internal primers and the ABI Prism™ dRhodamine Terminator Cycle Sequencing Kit (PE Applied Biosystems). Two subclones representing each type of *KIT* sequence were initially sequenced and in those cases where a discrepancy was observed (possibly due to PCR errors) additional clones were sequenced over those particular nucleotide sites. RT-PCR analysis of *KIT* exon 16-19 was carried out with the primers KIT1F (5'-TCR TAC ATA GAA AGA GAY GTG ACT C) and KIT7R (5'-AGC CTT CCT TGA TCA TCT TGT AG).

Results

The sequence of the KIT gene coding region derived from an animal of the Hampshire Breed is shown in Figure 10. Differences between *KIT* cDNA sequences cloned from a Hampshire and a Yorkshire/Landrace pig, respectively are shown in the table below. The sequence comparison includes the whole open reading frame 2919 bp, except for the last 27 bp occupied by the reverse PCR primer. Exon and base pair position number as well as amino acid codon are given for each difference. Polymorphic bases are shown in bold. A dash indicates identity with the Hampshire (*i*) allele.

Breed	Coat Colour	Assumed Genotype	Sequence variant	Splicing exon 14	Splicing exon 17	Exon 5 821	Exon 5 828	Exon 6 978	Exon 6 984	Exon 6 1008	Exon 9 1464	Exon 18 2502	Exon 19 2678
Hampshire	coloured	i/i	KIT1*0101	normal	normal	AGG Arg	ACA Thr	AAC Asn	GGA Gly	GAG Glu	ACG Thr	CCT Pro	GCG Ala
Yorkshire / Landrace	white	I/I[1]	KIT1*0201	normal	normal	-	-	AAT Asn	GGC Gly	GAA Glu	ACA Thr	CCC Pro	GTG Val
			KIT1*0202	normal	normal	AAG Arg	ACG Thr	-	-	GAA Glu	ACA Thr	CCC Pro	GTG Val
			KIT2*0101)	normal	skipped	-	-	AAT Asn	GGC Gly	GAA Glu	ACA Thr	CCC Pro	GTG Val
			KIT2 splice variant	skipped[2]	skipped	-	-	AAT Asn	GGC Gly	GAA Glu	ACA Thr	CCC [2]	GTG [2]

[1] Genotype I/I, I/I* or I*/I* inferred by the pig being a sow that got a 100 % white farrow following mating to a Hampshire (i/i) boar.

[2] The skipping of exon 14 (151 bp) cause a nonsense translation with termination at position 2161.

Example 17

DNA preparation

Genomic DNA was prepared as described in example 12.

PCR

A 158bp fragemnt covering 99bp of the end of exon 19 and 59bp of the KIT gene was amplified using forward primer LA93 (5'-GAG CAG CCC CTA CCC CGG AAT GCC AGT TGA-3') and reverse primer KIT56 (5'-CTT TAA AAC AGA ACA TAA AAG CGG AAA CAT CAT GCG AAG G-3'). PCR was carried out on a Perkin Elmer 9600 Thermal Cycler in a total volume of 20μl containing 25ng genomic DNA, 1.5mM $MgCl_2$, 50mM Kcl, 10mM Tris-HCl, pH 8.3, 200μM dNTPs, 0.5u AmpliTaq Gold (Perkin Elmer) and 10 pmol of both LA93 and KIT56 primer. To activate AmpliTaq Gold, initial heat denaturation was carried out at 94°C for 10 minutes followed by 32 cycles each consisting of 45 sec at 94°C, 45 sec at 55°C and 45 sec at 72°C.

Restriction digestion and electrophoresis

The PCR amplification product is 158 bp in length. To test for polymorphism at position 93 of this product (corresponding to position 2678 of the KIT cDNA sequence) digestion reactions were set up and incubated as below:

6.0μl PCR product
1.0μl 10x reaction buffer 3 (New England Biolabs)
0.2μl *Aci*I (5u/μl)
2.8μl deionised water

Following digestion at 37°C for 120 minutes each 10μl reaction volume had 2μl of loading dye aded and the mix was loaded on an 8% native polyacrylamide gel (Protogel, 37.5:1 acrylamide:bisacrylamide, National Diagnostics, Atlanta) in 0.5 x TBE (44.5 mM Tris pH8.0, 44.5 mM boric acid and 0.5 mM EDTA) and electrophoresed for 3 hours at 200v in a vertical slab gel unit (SE600 Hoefer Scientific Instruments). Products are visualised by ethidium bromide staining.

Results

The reverse primer is designed such that an *Aci*I site is introduced into the amplified sequence. This results in digestion of amplicon with AciI releasing a fragment of 23bp that allows confirmation of the digestion process. Digestion of the remaining 135 bp fragment into fragments of 92 and 43 bp is

dependant on the nucleotide at the position corresponding to position 2678 of the KIT cDNA sequence. T at this position prevents digestion while a C at this position allows digestion. Gel resolution is not sufficient to allow resolution of the 23bp fragment but comparison to undigested product allows confirmation of the process.

Figure 11 illustrates the results obtained with animals of a range of genotypes.

The test was used to analyse a total of 66 unrelated individuals from seven breeds of pig. The results are shown in the table below:

Breed	No.	KIT Genotype[1]	Genotype at pos'n 2678		
			C/C	C/T	T/T
Hampshire	4	*i/i*	1	1	2
Polish Wild Boar	13	*i/i*	0	1	12
Duroc	11	*i/i*	0	1	10
Pietrain	1	*i/i*	0	0	1
Swedish Wild Boar	1	*i/i*	1	0	0
Swedish Landrace	12	*I/I*	0	0	12
	5	*I/I*^{P}	0	2	3
Swedish Yorkshire	14	*I/I*	0	1	13
	5	*I/I*^{P}	0	1	4

1 Genotype based on *Nla*III RFLP analysis as described in example 11.

Example 18

Determination of genotype at the *I* locus using a rapid DNA based test

Crude DNA lysates were prepared from hair samples from animals of three breeding lines, a Hampshire based line, a Large White line, and white animals from a cross bred line originally produced from the two former lines. Four hair follicles were placed into 100μl of K buffer (50mM KCl, 10mM Tris-HCl pH 8.3, 2.5mM $MgCl_2$, 0.5% w/v Tween 20) and 1μl Proteinase K (15 mg/ml) (Boehringer) added. This mix was incubated for 2 hours at 55°C followed by 16 min at 95°C. DNA was also prepared as described in example 12.

Allelic discrimination reactions were set up using the PE Applied Biosystems TaqMan™ system. 25μl reactions contained the primers E19FOR (5-GAGCAGCCCCTACCCCGGAATGCCAGTTGA-3') and E19REV (5-CTTTAAAACAGAACATAAAAGCGGAAACATCATGCGAAGG-3') at 300nM, 8% glycerol (w/v) 1X TaqMan™ buffer A (PE Applied Biosystems), 5mM $MgCl_2$, 200μM dATP, dGTP, dCTP and dUTP, 0.65 units AmpliTaq Gold™ (PE Applied Biosystems), 0.25 units AmpErase™ UNG (PE Applied Biosystems) and the TaqMan™ probes E19PC (5'- CATACATTTCCGCAGGTGCATGC-FAM) and E19PT (5'- TCATACATTTCCACAGGTGCATGC-TET) at a concentration of 100mM. 1μl of crude lysate DNA was used as template. PCR amplification was carried out using a PE9600 thermal cycler (PE Applied Biosystems) or a the ABI7700 Prism (PE Applied Biosystems) with a thermal cycling regime of 50°C for 2 min followed by 95°C for 10 min followed by 40 cycles of 95°C 15 sec, 62°C 1 min. 8 control samples of each homozygote genotype, 2678C and 2678T, and 8 no template controls where deionized water was substituted for template controls were used per 96 well plate. Allele identification based on these reactions was carried out using the allelic discrimination function of the ABI7700 Prism (PE Applied Biosystems).

Results

The test was used to analyse a total of 20 unrelated individuals from four breeds of pig. The results are shown in the table below:

Breed	No.	Assumed KIT Genotype	Genotype at pos'n 2678		
			C/C	C/T	T/T
Hampshire	5	*i/i*	1	1	3
Landrace	5	*I/I*	0	0	5
Duroc	5	*i/i*	0	0	5
Pietrain	1	*i/i*	0	0	5

CLAIMS:

1. A method for:

(a) differentiating animal products on the basis of breed origin; or

(b) determining or testing the breed origin of an animal product; or

(c) validating an animal product;

comprising the steps of:

(i) providing a sample of the animal product; and

(ii) analysing the allele(s) of one or more breed determinant genes present in the sample.

2. The method of claim 1 wherein the breed determinant is a monogenic trait.

3. The method of claim 1 wherein the breed determinant is a polygenic trait.

4. The method of any one of claims 1-3 wherein the overt phenotypic trait is a behavioural or morphological trait.

5. The method of claim 3 or claim 4 wherein the overt phenotypic trait varies qualitatively or quantitatively between breeds.

6. The method of any one of the preceding claims wherein the breed determinant gene analysed in step (ii) is selected from any of:

(a) a coat colour gene; and/or

(b) a coat pattern gene; and/or

(c) a coat texture gene; and/or

(d) a coat density gene; and/or

(e) a coat length gene; and/or

(f) a ear aspect gene; and/or

(g) a double muscling gene; and/or

(h) a horn morphology gene; and/or

(i) a tusk morphology gene; and/or

(j) an eye colour gene; and/or

(k) a plumage gene; and/or

(l) a beak colour/morphology gene; and/or

(m) a vocalization (e.g. barking) gene; and/or

(n) a comb or wattle gene; and/or

(o) a gene controlling display behaviour.

7. The method of claim 6(a) wherein the coat colour gene is the KIT or αMSHR gene (for example, the pig KIT or αMSHR gene).

8. The method of any one of the preceding claims wherein the sample is a nucleic acid sample and the analysing step (ii) comprises DNA or RNA analysis.

9. The method of any one of claims 1-7 wherein the sample is a protein sample and the analysing step (ii) comprises protein analysis.

10. The method of claim 8 wherein the analysis step (ii) comprises:

(a) selectively amplifying a specific fragment of nucleic acid (e.g. by PCR); and/or

(b) testing for the presence of one or more restriction endonuclease sites within the breed determinant gene(s) (e.g. restriction fragment length polymorphism (RFLP) analysis); and/or

(c) determining the nucleotide sequence of all or a portion of the breed determinant gene(s); and/or

(d) probing the nucleic acid sample with an allele-specific DNA or RNA probe: and/or

(e) carrying out one or more PCR amplification cycles of the nucleic acid sample using at least one pair of suitable primers and then carrying out RFLP analysis on the amplified nucleic acid so obtained.

11. The method of claim 9 wherein the analysis step (ii) comprises:

(a) probing the protein sample with an antibody (e.g. a monoclonal antibody) specific for an allele-specific epitope; and/or

(b) electrophoretic analysis; and/or

(c) chromatographic analysis; and/or

(d) amino-acid sequence analysis; and/or

(e) proteolytic cleavage analysis; and/or

(f) epitope mapping.

12. The method of claim 7 wherein the analysis step (ii) comprises determining the nucleotide sequence of the KIT or αMSHR gene or the amino acid sequence of the KIT or αMSHR protein.

13. The method of claim 7 or claim 12 wherein the analysis step (ii) comprises establishing the presence or absence of at least one nucleotide change in the coding region of the KIT or αMSHR gene.

14. The method of any one of claims 7, 12 and 13 wherein the analysis step (ii) further comprises determining the association between one or more microsatellite marker alleles linked to the KIT or αMSHR gene and to particular alleles of the KIT or αMSHR gene.

15. The method of claim 14 wherein the analysis step (ii) is based on the identification of microsatellite markers present in the nucleic acid sample.

16. The method of claim 7 wherein the analysis step (ii) comprises:

(a) determining the association between one or more microsatellite marker alleles linked to the KIT or αMSHR gene and to particular alleles of the KIT or αMSHR gene;

(b) determining which microsatellite marker allele or alleles are present in the nucleic acid sample.

17. The method of any one of claims 7 and 12-16 wherein the analysis step (ii) further comprises the step of determining the genotype of at least one additional locus.

18. The method of claim 17 wherein the additional locus is an additional coat colour locus.

19. The method of claim 18 wherein the additional coat colour locus is the *KIT* gene locus (e.g. the pig *KIT* gene locus).

20. The method of claim 17 wherein the additional locus is a breed determinant gene locus selected from any of those genes specified in claim 6.

21. The method of claim 17 wherein the additional locus is a breed specific marker.

22. The method of claim 21 wherein the breed specific marker is a microsatellite marker.

23. The method of any one of claims 7 and 12-22 wherein the analysis step (ii) comprises PCR using at least one pair of suitable primers.

24. The method of claim 23 wherein the gene is the pig αMSHR gene and the at least one pair of suitable primers is:

αMSHR Forward Primer 1: (5'-TGT AAA ACG ACG GCC AGT RGT GCC TGG AGG TGT CCA T-3');

αMSHR Forward Primer 5: (5'-CGC CCA GAT GGC CGC GAT GGA CCG-3')

αMSHR Forward Primer 2: (5'-CGG CCA TCT GGG CGG GCA GCG TGC-3')
αMSHR Reverse Primer 2: (5'-GGA AGG CGT AGA TGA GGG GGT CCA-3')
αMSHR Forward Primer 3: (5'-GCA CAT CGC CCG GCT CCA CAA GAC-3')
αMSHR Reverse Primer 3: (5'-GGG GCA GAG GAC GAC GAG GGA GAG-3')

25. The method of any one of claims 7 and 12-24 wherein the analysis step (ii) comprises restriction fragment length polymorphism (RFLP) analysis, for example involving digesting the pig nucleic acid with one or more of the restriction enzymes *Bst*UI, *Hha*I and/or *Bsp*HI.

26. The method of claim 25 wherein the gene is the pig αMSHR gene and the analysis involves identification of a polymorphism at nucleotide position 400, 422, 480, 487, 608, 844, 846 or in the region 482-487 or 841-849 of the coding sequence of the pig αMSHR gene.

27. The method of claim 7 wherein the analysis step (ii) carrying out one or more PCR amplification cycles of the nucleic acid sample using at least one pair of suitable primers and then carrying out RFLP analysis on the amplified nucleic acid so obtained to determine the KIT or αMSHR genotype of the pig.

28. The method of claim 27 wherein the gene is the pig αMSHR gene and the at least one pair of suitable primers is as defined in claim 24.

29. The method of claim 27 or 28 wherein the gene is the pig KIT or αMSHR gene and the RFLP analysis is as defined in claim 26.

30. The method of any one of the preceding claims wherein the animal product is meat (e.g. processed and/or canned meat), egg, egg swab or washing, semen, wool or leather.

31. The method of any one of the preceding claims wherein the sample comprises genomic DNA, RNA or mitochondrial DNA.

32. The method of any one of the preceding claims wherein the animal is mammal (e.g. pig, cattle, dog, cat, horse, sheep, rodent or rabbit), fish (e.g. salmon or trout) or bird (chicken or turkey).

33. A kit for:

(a) differentiating animal products on the basis of breed origin; or

(b) determining or testing the breed origin of an animal product; or

(c) validating an animal product;

comprising one or more reagents for analysing the allele(s) of one or more breed determinant genes present in the sample.

Figure 1

a: Nucleotide sequence

```
                                                                                                                294
Wildboar    ??C CCG AAC GGG CTC TTC CTC AGC CTG GGG CTG GTG AGC CTC GTG GAG AAC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                345
Wildboar    GTG CTG GTG GTG GCC GCC ATC GCC AAG AAC CGC AAC CTG CAC TCG CCC ATG
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                396
Wildboar    TAC TAC TTC GTC TGC TGC CTG GCC GTG TCG GAC CTG CTG GTG AGC GTG AGC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                447
Wildboar    AAC GTG CTG GAG ACG GCC GTG CTG CTG CTG CTG GAG GCG GGC GCC CTG GCC
Meishan     ... A.. ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                498
Wildboar    GCC CAG GCC GCC GTG GTG CAG CAG CTG GAC AAT GTC ATG GAC GTG CTC ATC
Meishan     ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... A.. ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... A.. ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                549
Wildboar    TGC GGC TCC ATG GTG TCC AGC CTC TGC TTC CTG GGC GCC ATC GCC GTG GAC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

600
Wildboar    CGC TAC GTG TCC ATC TTC TAC GCG CTG CGC TAC CAC AGC ATC GTG ACG CTG
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

Figure 1 continued

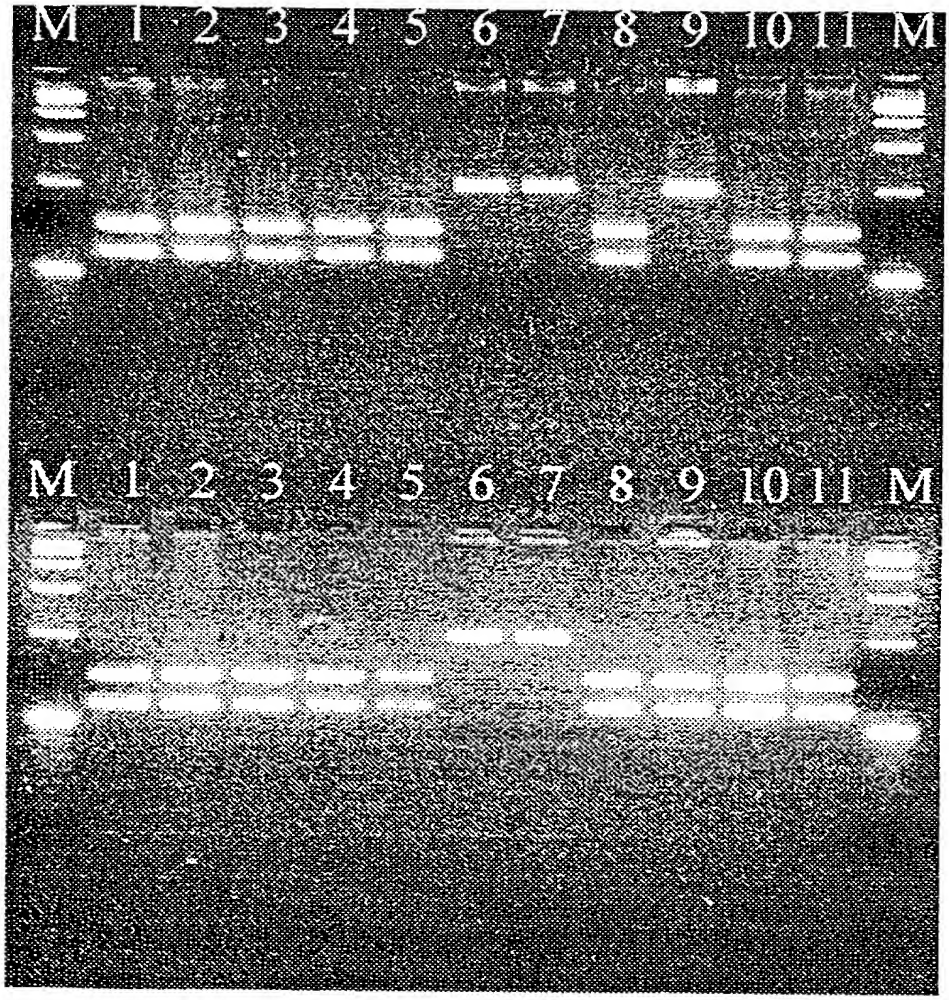
```
                                                                                                                651
Wildboar    CCC CGC GCG GGG CGG GCT ATC GCG GCG ATC TGG GCG GGC AGC GTG CTC TCC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... .T. ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                702
Wildboar    AGC ACC CTC TTC ATC GCC TAC TAC CAC CAC ACG GCC GTC CTG CTG GGC CTC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                753
Wildboar    GTC AGC TTC TTC GTG GCC ATG CTG GCG CTC ATG GCG GTA CTG TAC GTC CAC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                804
Wildboar    ATG CTG GCC CGG GCC TGC CAG CAC GGC CGG CAC ATC GCC CGG CTC CAC AAG
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

                                                                                                                855
Wildboar    ACG CAG CAC CCC ACC CGC CAG GGC TGC GGC CTC AAG GGC GCG GCC ACC CTC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... A.. ... ... ...

                                                                                                                906
Wildboar    ACC ATC CTG CTG GGC GTC TTC CTC CTC TGC TGG GCA CCC TTC TTC CTG CAC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

957
Wildboar    CTC TCC CTC GTC GTC CTC TGC CCC CAG CAC CCC ACC TGC GGC TGC GTC TTC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...

Wildboar    AAG AAC GTC AAC CTC TTT CTG GCC CTC GTC ATC TGC AAC TCC ATC
Meishan     ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Largewhite  ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Hampshire   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Duroc       ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
```

Figure 1 continued

b. Amino acid sequence

```
                                                                                                    99
Wildboar    ?PNGLFLSLG LVSLVENVLV VAAIAKNRNL HSPMYYFVCC LAVSDLLVSV SNVLETAVLL
Meishan     .......... .......... .......... .......... .......... ..M.....P
Largewhite  .......... .......... .......... .......... .......... ..........
Hampshire   .......... .......... .......... .......... .......... ..........
Duroc       .......... .......... .......... .......... .......... ..........

                                                                                                   159
Wildboar    LLEAGALAAQ AAVVQQLDNV MDVLICGSMV SSLCFLGAIA VDRYVSIFYA LRYHSIVTLP
Meishan     .......... .......... .......... .......... .......... ..........
Largewhite  .......... .......... .N........ .......... .......... ..........
Hampshire   .......... .......... .N........ .......... .......... ..........
Duroc       .......... .......... .......... .......... .......... ..........

                                                                                                   219
Wildboar    RAGRAIAAIW AGSVLSSTLF IAYYHHTAVL LGLVSFFVAM LALMAVLYVH MLARACQHGR
Meishan     .......... .......... .......... .......... .......... ..........
Largewhite  .......... .......... .......... .......... .......... ..........
Hampshire   .......... .......... .......... .......... .......... ..........
Duroc       .V........ .......... .......... .......... .......... ..........

                                                                                                   279
Wildboar    HIARLHKTQH PTRQGCGLKG AATLTILLGV FLLCWAPFFL HLSLVVLCPQ HPTCGCVFKN
Meishan     .......... .......... .......... .......... .......... ..........
Largewhite  .......... .......... .......... .......... .......... ..........
Hampshire   .......... .......... .......... .......... .......... ..........
Duroc       .......... .......... T......... .......... .......... ..........

Wildboar    VNLFLALVIC NSI
Meishan     .......... ...
Largewhite  .......... ...
Hampshire   .......... ...
Duroc       .......... ...
```

Figure 2



*Bst*UI

*Hha*I

Figure 3

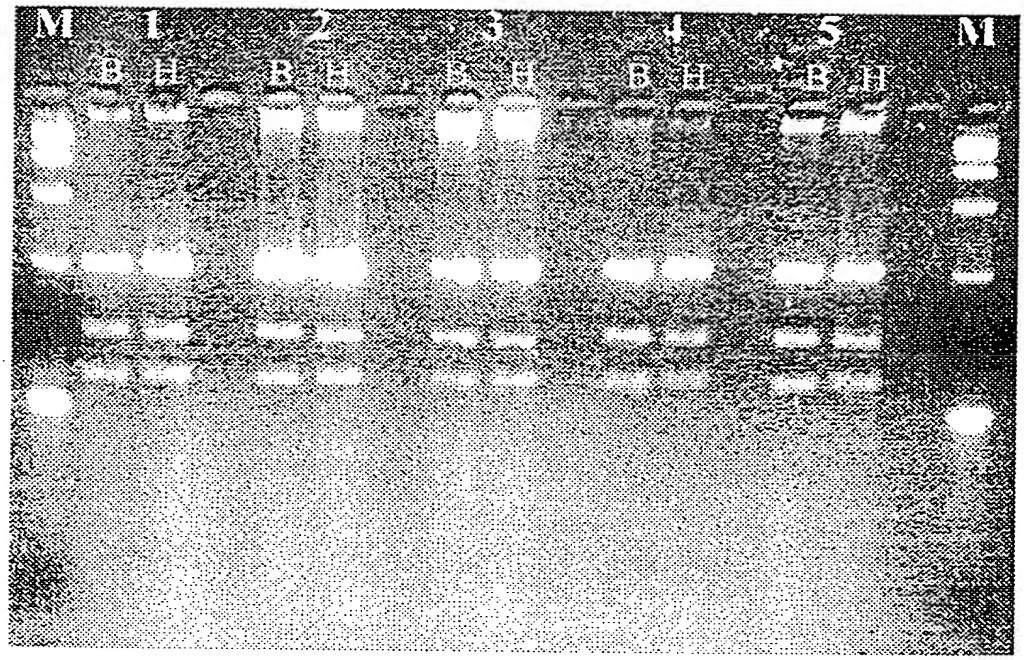


Fig 4

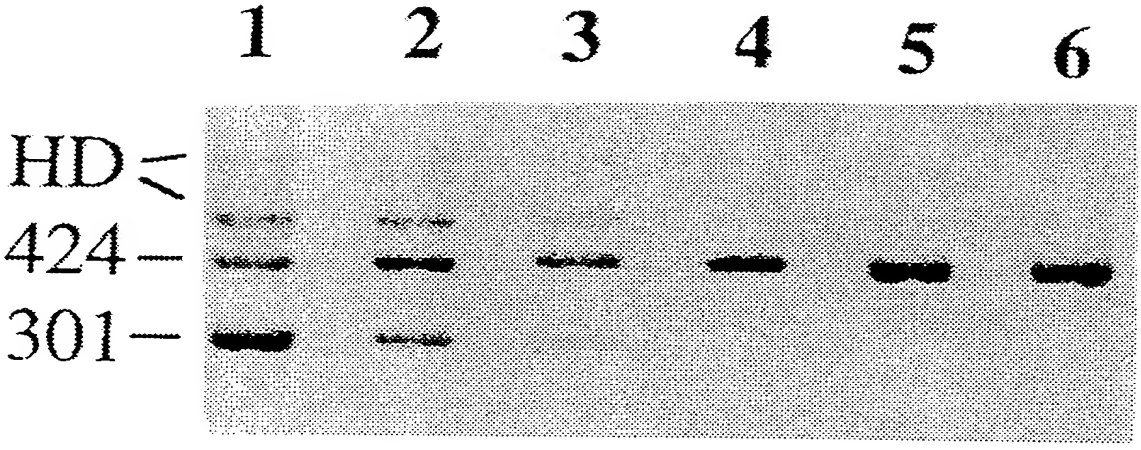
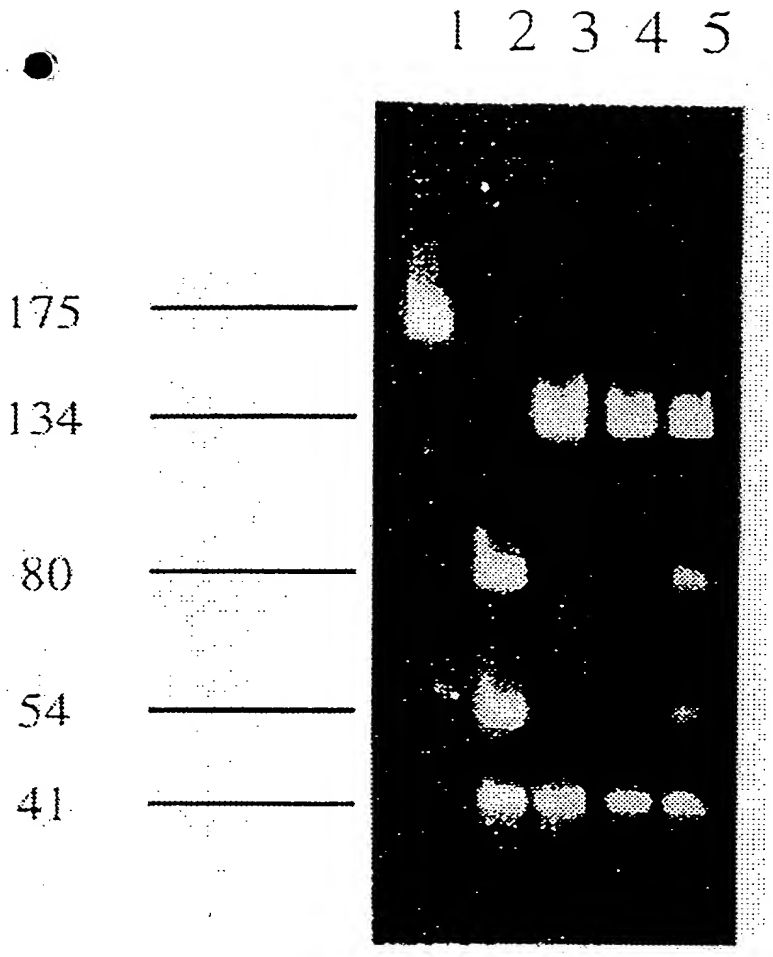


Figure 5

Sequence Alignment Across the Exon /Intron Border of KIT Exon 17

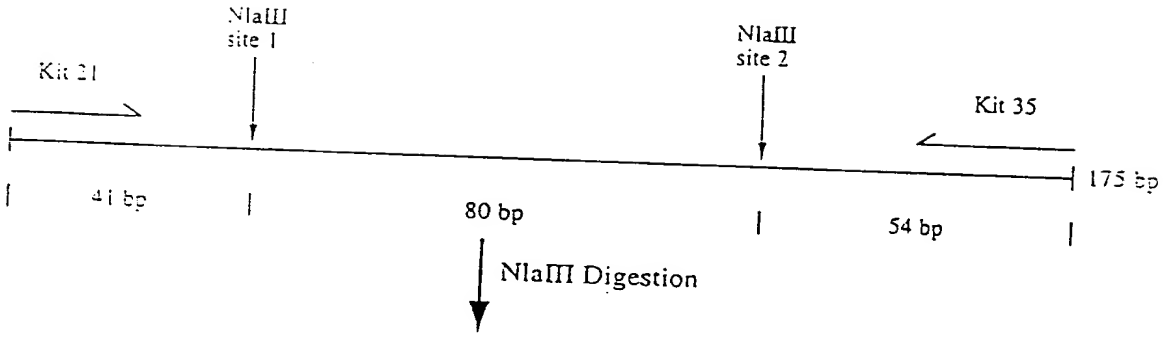
Allele	Gene Copy	Sequence (Exon 17 | Intron 17 ↓)
I	KIT1	AAT TAC GTG GTC AAA GGA AAC|GTG AGT ACC CAC GCT CTC CTG ACA GTC
	KIT2	... ...|A.. ... ...
I^P	KIT1	... ...|G... ... ...
	KIT1	... ...|G.. ... ...
i	KIT1	... ...|G.. ... ...

8/14


1 2 3 4 5
175
134
80
54
41

6C

6 A


NlaIII site 1
NlaIII site 2
Kit 21
Kit 35
175 bp
41 bp
80 bp
54 bp
NlaIII Digestion

6B

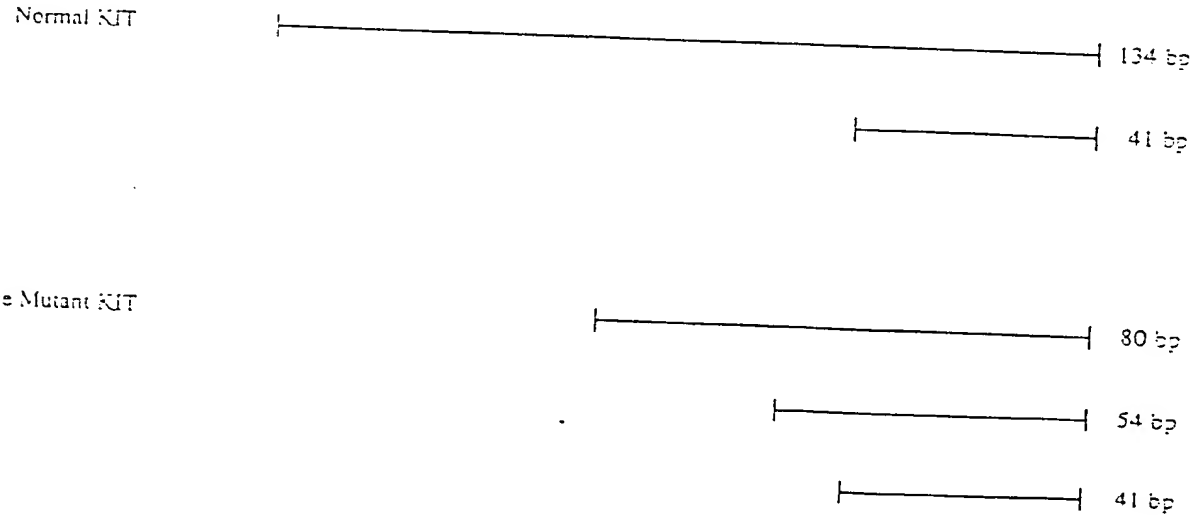

Normal KIT
134 bp
41 bp
Splice Mutant KIT
80 bp
54 bp
41 bp

Figure 7

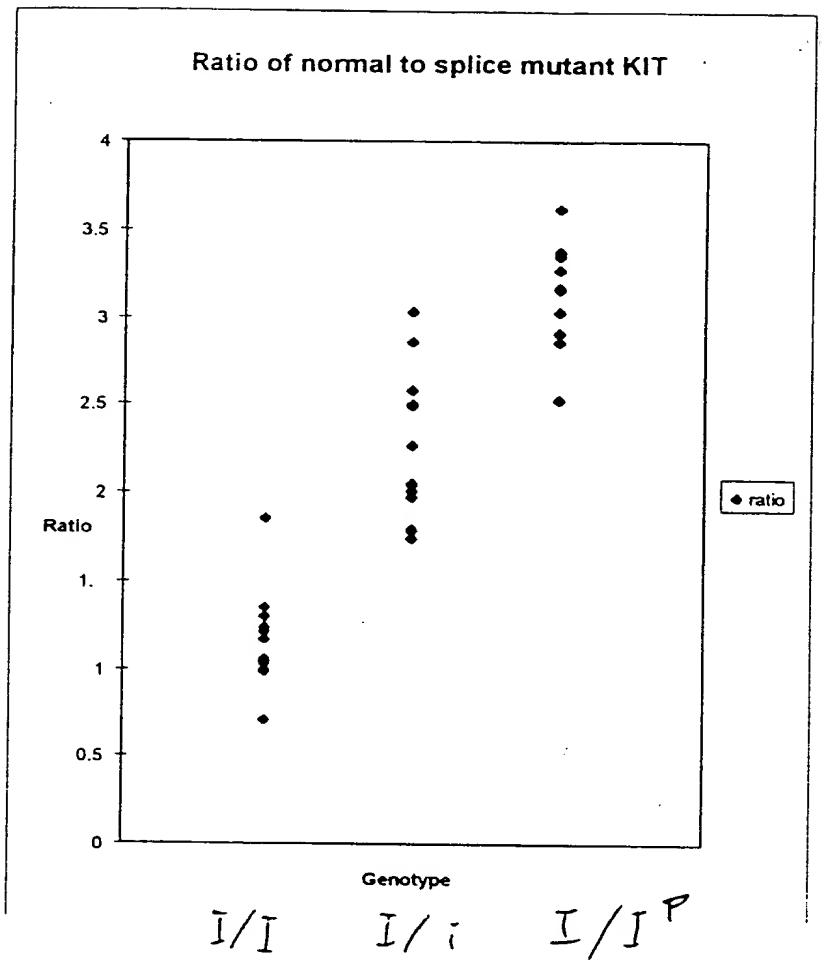
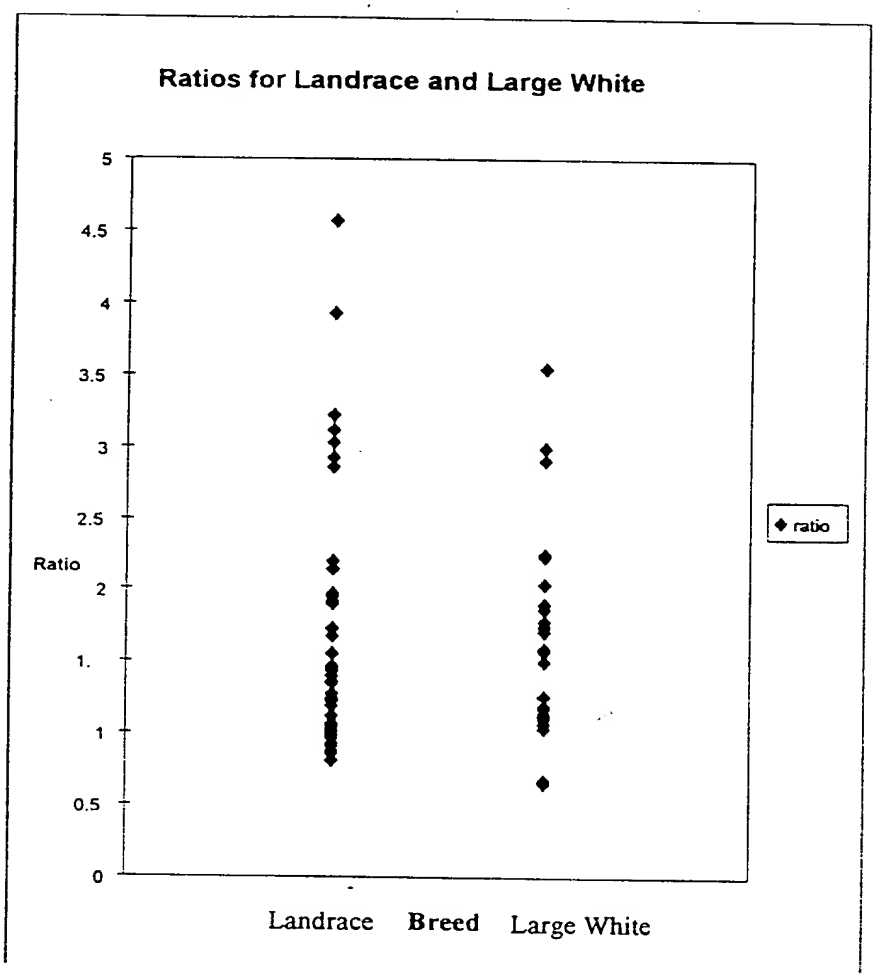
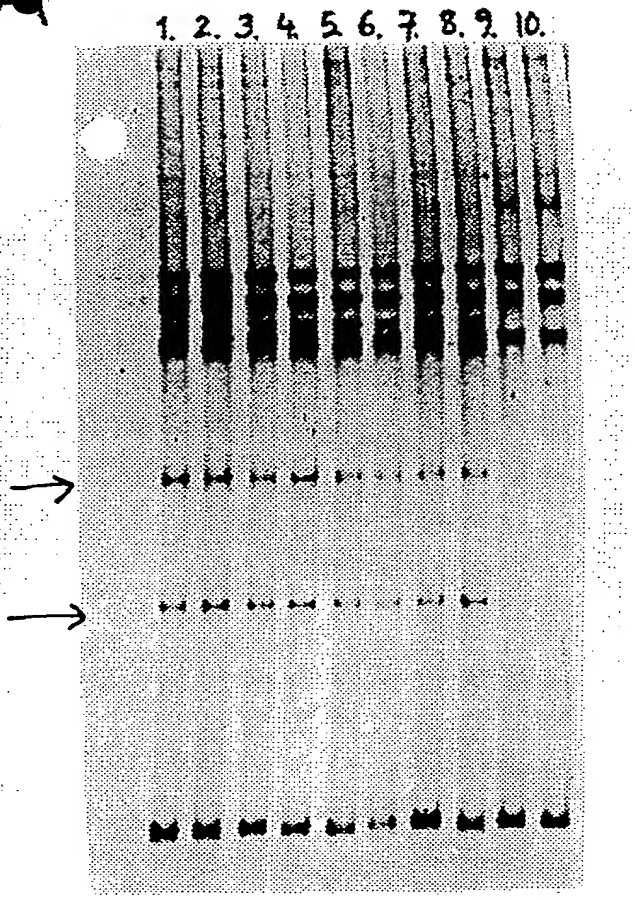

Ratio of normal to splice mutant KIT
4
3.5
3
2.5
2
1.
1
0.5
0
Ratio
Genotype
ratio
I/I
I/i
I/I^P

Figure 8


Ratios for Landrace and Large White
5
4.5
4
3.5
3
2.5
Ratio
2
1.
1
0.5
0
ratio
Landrace
Breed
Large White

A

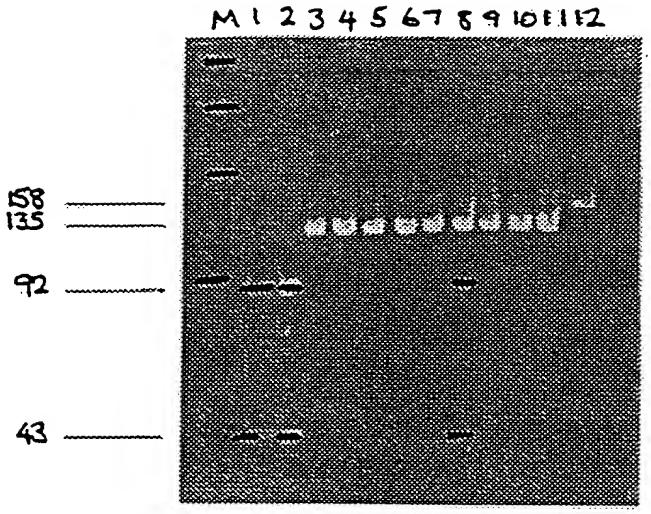


11/14 Fig 9

Figure 10

1
ATG AGA GGC GCT CGC CGC GCC TGG GAT TTT CTC TTC GTC CTG CAG CTC TTG
52
CTT CGC GTC CAG ACA GGC TCT TCT CAG CCA TCT GTG AGT CCA GAG GAA CTG
103
TCT CCA CCA TCC ATC CAT CCA GCA AAA TCA GAG TTA ATC GTC AGT GCT GGC
154
GAT GAG ATT AGG CTG TTC TGC ACC GAT CCA GGA TCT GTC AAA TGG ACT TTT
205
GAG ACC CTG GGT CAG CTG AGT GAG AAT ACA CAC GCA GAG TGG ATC GTG GAG
256
AAA GCA GAG GCC ATG AAT ACA GGC AAT TAT ACA TGC ACC AAT GAA GGC GGT
307
TTA AGC AGT TCC ATT TAT GTG TTT GTT AGA GAT CCT GAG AAG CTT TTC CTC
358
GTC GAC CCT CCC TTG TAT GGG AAG GAG GAC AAT GAC GCG CTG GTC CGA TGT
409
CCT CTG ACG GAC CCA GAG GTG ACC AAT TAC TCC CTC ACG GGC TGC GAG GGG
460
AAA CCC CTT CCC AAG GAT TTG ACC TTC GTC GCG GAC CCC AAG GCC GGC ATC
511
ACC ATC AGA AAC GTG AAG CGC GAG TAT CAT CGG CTC TGT CTC CAC TGC TCC
562
GCC AAC CAG GGG GGC AAG TCC GTG CTG TCG AAG AAA TTC ACC CTG AAA GTG
613
AGG GCA GCC ATC AGA GCT GTA CCT GTT GTG GCT GTG TCC AAA GCA AGC TAC
664
CTT CTC AGG GAA GGG GAG GAA TTT GCC GTG ATG TGC TTG ATC AAA GAC GTG
715
TCT AGT TCC GTG GAC TCC ATG TGG ATC AGG GAG AAC AGC CAG ACT AAA GCA
766
CAG GTG AAG AGG AAT AGC TGG CAT CAG GGT GAC TTC AAT TTT CTG CGG CAG
817
GAA AGG CTG ACA ATC AGC TCA GCA AGA GTT AAT GAT TCT GGC GTG TTC ATG
868
TGT TAC GCC AAT AAT ACT TTT GGA TCT GCA AAT GTC ACA ACC ACC TTA GAA
919
GTA GTA GAT AAA GGA TTC ATT AAT ATC TTC CCT ATG ATG AAT ACC ACT GTG
970
TTT GTA AAC GAT GGA GAG GAT GTG GAT CTA ATT GTT GAG TAC GAG GCG TAC
1021
CCC AAA CCT GAA CAC CGA CAG TGG ATA TAT ATG AAC CGC ACT GCC ACT GAT
1072
AAG TGG GAG GAT TAT CCC AAG TCT GAG AAT GAA AGT AAC ATC AGA TAT GTA
1123
AGT GAA CTT CAC TTG ACC AGA TTA AAA GGG ACC GAA GGA GGC ACT TAC ACA
1174
TTT CTC GTG TCC AAT GCT GAT GTC AAT TCT TCT GTG ACA TTT AAT GTT TAC
1225
GTG AAC ACA AAA CCA GAA ATC CTG ACT CAT GAC AGG CTC ATG AAC GGC ATG
1276
CTC CAG TGT GTG GCG GCA GGC TTC CCA GAG CCC ACC ATC GAT TGG TAT TTC
1327
TGT CCA GGC ACC GAG CAG AGA TGT TCC GTT CCC GTT GGG CCA GTG GAC GTG
1378
CAG ATC CAA AAC TCA TCT GTA TCA CCG TTT GGA AAA CTA GTG ATT CAC AGC
1429
TCC ATT GAT TAC AGT GCA TTC AAA CAC AAC GGC ACG GTG GAG TGC AGG GCT

1480
TAC AAC GAT GTG GGC AAG AGT TCT GCC TTT TTT AAC TTT GCA TTT AAA GAA
1531
CAA ATC CAT GCC CAC ACC CTC TTC ACG CCT TTG CTG ATT GGT TTT GTG ATC
1582
GCA GCG GGT ATG ATG TGT ATC ATC GTG ATG ATT CTC ACC TAT AAA TAT CTA
1633
CAG AAG CCC ATG TAT GAA GTA CAG TGG AAG GTT GTC GAG GAG ATA AAT GGA
1684
AAC AAT TAT GTC TAC ATA GAC CCA ACG CAA CTT CCT TAT GAT CAC AAA TGG
1735
GAA TTT CCC AGG AAC AGG CTG AGT TTT GGC AAA ACC TTG GGT GCT GGC GCC
1786
TTC GGG AAA GTC GTT GAG GCC ACT GCA TAC GGC TTA ATT AAG TCA GAT GCG
1837
GCC ATG ACC GTT GCC GTG AAG ATG CTC AAA CCA AGT GCC CAT TTA ACG GAA
1888
CGA GAA GCC CTA ATG TCT GAA CTC AAA GTC TTA AGT TAC CTC GGT AAT CAC
1939
ATG AAT ATT GTG AAT CTT CTC GGC GCC TGC ACC ATT GGA GGG CCC ACC CTG
1990
GTC ATT ACA GAA TAT TGT TGC TAT GGT GAT CTC CTG AAT TTT TTG AGA CGG
2041
AAA CGT GAT TCG TTT ATT TGC TCA AAG CAG GAA GAT CAC GCA GAA GCG GCG
2092
CTT TAT AAG AAC CTT CTG CAT TCA AAG GAG TCT TCC TGC AGT GAC AGT ACT
2143
AAC GAG TAC ATG GAC ATG AAA CCC GGA GTG TCT TAT GTG GTA CCA ACC AAG
2194
GCA GAC AAA AGG AGA TCT GCG AGA ATA GGC TCA TAC ATA GAA CGA GAT GTG
2245
ACT CCT GCC ATC ATG GAA GAT GAT GAG TTG GCC CTA GAC CTG GAG GAC TTG
2296
CTC AGC TTT TCT TAC CAA GTG GCA AAG GGC ATG GCC TTC CTC GCC TCG AAG
2347
AAT TGT ATT CAC AGA GAC TTG GCG GCC AGA AAT ATC CTC CTT ACT CAT GGT
2398
CGA ATC ACA AAG ATT TGT GAT TTT GGT CTA GCC AGA GAC ATC AAG AAT GAT
2449
TCT AAT TAC GTG GTC AAA GGA AAC GCT CGG CTA CCC GTG AAG TGG ATG GCA
2500
CCT GAG AGC ATT TTC AAC TGT GTC TAC ACA TTT GAA AGC GAT GTC TGG TCC
2551
TAT GGG ATT TTT CTG TGG GAG CTC TTC TCT TTA GGG AGC AGC CCC TAC CCC
2602
GGA ATG CCA GTT GAT TCT AAA TTC TAC AAG ATG ATC AAG GAG GGT TTC CGA
2653
ATG CTC AGC CCT GAG CAT GCA CCT GCG GAA ATG TAT GAC ATC ATG AAG ACT
2704
TGC TGG GAT GCG GAT CCC CTC AAA AGA CCA ACG TTT AAG CAG ATC GTG CAG
2755
CTG ATT GAG AAG CAG ATT TCG GAG AGC ACC AAT CAC ATT TAT TCC AAC TTA
2806
GCG AAC TGC AGC CCC CAC CGG GAG AAC CCC GCG GTG GAT CAT TCT GTG CGG
2857
ATC AAC TCC GTG GGC AGC AGT GCC TCC TCC ACG CAG CCT CTG CTT GTC CAC
2908
GAA GAT GTC TGA


M 1 2 3 4 5 6 7 8 9 10 11 12
158
135
92
43